

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
7 September 2001 (07.09.2001)

PCT

(10) International Publication Number  
WO 01/64710 A2

(51) International Patent Classification<sup>7</sup>:

C07K

(72) Inventors; and

(21) International Application Number: PCT/US01/06699

(75) Inventors/Applicants (for US only): DRAGIC, Tatjana [US/US]; 191 Fox Meadow Road, Scarsdale, NY 10583 (US). OLSON, William, C. [US/US]; 21 Fawn Court, Ossining, NY 10562 (US).

(22) International Filing Date: 28 February 2001 (28.02.2001)

(74) Agent: WHIITE, John, P.; Cooper & Dunham LLP, 1185 Avenue of the Americas, New York, NY 10036 (US).

(25) Filing Language:

English

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(26) Publication Language:

English

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(30) Priority Data:

|            |                               |    |
|------------|-------------------------------|----|
| 60/185,667 | 29 February 2000 (29.02.2000) | US |
| 60/205,839 | 19 May 2000 (19.05.2000)      | US |
| 60/267,231 | 7 February 2001 (07.02.2001)  | US |

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

|          |                               |
|----------|-------------------------------|
| US       | 60/185,667 (CIP)              |
| Filed on | 29 February 2000 (29.02.2000) |
| US       | 60/205,839 (CIP)              |
| Filed on | 19 May 2000 (19.05.2000)      |
| US       | 60/267,231 (CIP)              |
| Filed on | 7 February 2001 (07.02.2001)  |

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(71) Applicants (for all designated States except US): PROGENICS PHARMACEUTICALS, INC. [US/US]; 777 Old Saw Mill River Road, Tarrytown, NY 10591 (US). AARON DIAMOND AIDS RESEARCH CENTRE [US/US]; 7th floor, 455 First Avenue, New York, NY 10016 (US).

A2

(54) Title: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION

WO 01/64710

(57) Abstract: This invention provides a compound comprising the structure:  $\theta\alpha\text{YDINYYTSE}\beta\lambda$  wherein each T represents a threonine, each s represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine; wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction; wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction; wherein  $\theta$  represents an amino group or an acetylated amino group; wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group; wherein all of  $\alpha$ , Y, D, I, N, Y, Y, T, S, E and  $\beta$  are joined together by peptide bonds; further provided that at least two tyrosines in the compound are sulfated.

Applicants: Olson and Maddon  
Serial No. : 09/828/615  
Filed: April 6, 2001  
Exhibit 32

SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION

5 This application is a continuation-in-part of and claims  
the benefit of U.S. Provisional Application No. 60/267,231,  
filed February 7, 2001, U.S. Provisional Application No.  
60/205,839, filed May 19, 2000 and U.S. Provisional  
Application No. 60/185,667, filed February 29, 2000, the  
10 contents of which are hereby incorporated by reference into  
this application.

The invention disclosed herein was made with Government  
support under NIH Grant Nos. R01AI43847 (T.D.) and  
15 R01DK54718 (T.P.S.) from the Department of Health and Human  
Services. Accordingly, the government has certain rights in  
this invention.

Throughout this application, various publications are  
20 referenced within parentheses. Disclosures of these  
publications in their entireties are hereby incorporated by  
reference into this application to more fully describe the  
state of the art to which this invention pertains. Full  
bibliographic citations for these references may be found  
25 immediately preceding the claims.

Background of the Invention

HIV-1 entry into target cells is mediated by the successive  
interaction of the envelope glycoprotein gp120 with CD4 and  
30 a co-receptor belonging to the seven trans-membrane G

-2-

protein-coupled chemokine receptor family (Berger *et al.* Ann. Rev. Immunol. 17:657, 1999). Binding of gp120 to CD4 exposes or creates a co-receptor binding site on gp120 (Trkola *et al.* Nature 384:184, 1996, Wu *et al.* Nature, 5 384:179, 1996). CCR5 and CXCR4 are the most physiologically relevant and widely used HIV-1 co-receptors (Zhang and Moore, J. Virol. 73:3443, 1999). CCR5 mediates the entry of R5 isolates and CXCR4 mediates the entry of X4 isolates. R5X4 isolates are able to exploit both co-receptors (Berger 10 *et al.* Ann. Rev. Immunol. 17:657, 1999). It has been demonstrated that specific amino acids including acidic residues and tyrosines located within the CCR5 amino-terminal domain (Nt, amino acids 2-31) are essential for CCR5-mediated fusion and entry of R5 and R5X4 HIV-1 strains 15 (Dragic *et al.* J. Virol. 72:279, 1998; Rabut *et al.* J. Virol. 72:3464, 1998; Farzan *et al.* J. Virol. 72:1160, 1998; Doranz *et al.* J. Virol. 71:6305, 1997). More recently, Farzan *et al.* demonstrated that tyrosine residues in the CCR5 Nt are sulfated (Farzan *et al.* Cell 96:667, 20 1999)

Inhibition of cellular sulfation pathways, including tyrosine sulfation, by sodium chlorate decreased the 25 binding of a gp120/CD4 complex to CCR5<sup>+</sup> cells (Farzan *et al.* Cell 96:667, 1999). A number of prior reports had implicated a role for sulfate moieties in HIV-1 entry. Several sulfated compounds, such as dextran sulfate, can inhibit HIV-1 entry by associating with CD4 or gp120 30 (Baeuerle and Huttner J. Cell Biol 105:2655, 1987; Baba *et al.* Proc. Natl. Acad. Sci. USA 85:6132, 1998). Sulfated

proteoglycans have been shown to bind to HIV-1 gp120 at or near its third variable (V3) loop, which also determines co-receptor usage (Roderiquez et al. J. Virol. 69:2233, 1995; Hwang et al. Science 253:71, 1991). It is therefore 5 conceivable that sulfo-tyrosines in the CCR5 Nt also interact with gp120, increasing its affinity for CCR5. The reduction in gp120/CD4 binding caused by the pre-treatment of target cells with sodium chlorate, however, cannot be formally attributed to a reduction in CCR5 tyrosine 10 sulfation since chlorate can inhibit the sulfation of both tyrosines and proteoglycans.

The region of the CCR5 Nt spanning amino acids 2-18 contains residues that are critically important for viral 15 entry (Dragic et al. J. Virol. 72:279, 1998; Rabut et al. J. Virol. 72:3464, 1998; Farzan et al. J. Virol. 72:1160, 1998; Doranz et al. J. Virol. 71:6305, 1997). We previously demonstrated that tyrosines at positions 3, 10 and 14 were required for optimal co-receptor function, 20 whereas the Tyr15Phe substitution had little effect on entry (Rabut et al. J. Virol. 72:3464, 1998). Taken together, these findings suggested that HIV-1 entry may be critically dependent upon sulfation of Tyr-3, -10 and -14, but not Tyr-15. We therefore explored the role of sulfo- 25 tyrosines in positions 3, 10 and 14 by synthesizing peptides corresponding to amino acids 2-18 of the CCR5 Nt and carrying different tyrosine modifications. We first tested the ability of the Nt peptides to inhibit binding of gp120/CD4 complexes and anti-CCR5 Mabs to CCR5<sup>+</sup> cells. The 30 specific association of certain peptides with gp120/sCD4 complexes or with anti-CCR5 Mabs was further confirmed by

surface plasmon resonance (BIAcore) analysis. Inhibition of HIV-1 entry by the CCR5 Nt peptides was also tested. Our results suggest that amino acids 2-18 of the CCR5 Nt compose a gp120-binding site that determines the specificity of the interaction between CCR5 and gp120s from R5 and R5X4 isolates. Post-translational sulfation of the tyrosine residues in the CCR5 Nt is required for gp120 binding and may critically modulate the susceptibility of target cells to HIV-1 infection *in vivo*.

10

CCR5's normal physiologic activities involve binding and transducing signals mediated by CC-chemokines, including RANTES, MIP-1 $\alpha$  and MIP-1 $\beta$ , which direct activation and trafficking of T cells and other inflammatory cells. As such, CCR5 plays an important role in mediating the inflammatory reaction of diseases such as rheumatoid arthritis and multiple sclerosis. The synovial fluid of rheumatoid arthritis patients is highly enriched in CCR5-expressing T cells (Qin *et al.* J Clin Invest 101:746, 1998), and CCR5 is the predominant CC chemokine receptor expressed on T cells in the rheumatoid synovium (Gomez-Reino *et al.* Arthritis Rheum 42:989, 1999). Similarly, infiltration by CCR5-expressing cells is characteristic of plaque lesions in patients with multiple sclerosis (Balashov *et al.* Proc Natl Acad Sci USA 96:6873, 1999). Such observations provide a rationale for the use of agents that block CCR5 for therapy of inflammatory/autoimmune diseases, including but not limited to arthritis, multiple sclerosis, asthma, psoriasis, autoimmune diabetes, transplant rejection, and atherosclerosis.

Summary of the Invention

This invention provides a compound comprising the structure:



wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine; wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction; wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction; wherein  $\Theta$  represents an amino group or an acetylated amino group; wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group; wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined together by peptide bonds; further provided that at least two tyrosines in the compound are sulfated.

This invention also provides a compound comprising the structure:



wherein each T represents a threonine, each S represents a

serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine; wherein  $\alpha$  represents from 0 to 9 amino acids,  
5 with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction; wherein  $\beta$   
10 represents from 0 to 333 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom  
15 in the carboxy terminal direction;  
wherein  $\theta$  represents an amino group or an acetylated amino group; wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group; wherein all of  $\alpha$ , Y, D, I, N, Y, Y, T, S, E and  $\beta$  are joined together by peptide bonds; further provided that  
20 at least two tyrosines in the compound are sulfated.

This invention provides a composition which comprises a carrier and an amount of one of the compounds described herein effective to inhibit binding of HIV-1 to a CCR5 receptor on the surface of a CD4+ cell.  
25

This invention provides a method of inhibiting human immunodeficiency virus infection of a CD4+ cell which also carries a CCR5 receptor on its surface which comprises  
30 contacting the CD4+ cell with an amount of one of the

compounds described herein effective to inhibit binding of human immunodeficiency virus to the CCR5 receptor so as to thereby inhibit human immunodeficiency virus infection of the CD4+ cell.

This invention provides a method of preventing CD4+ cells of a subject from becoming infected with human immunodeficiency virus which comprises administering to the subject an amount of one of the compounds described herein 10 effective to inhibit binding of human immunodeficiency virus to CCR5 receptors on the surface of the CD4+ cells so as to thereby prevent the subject's CD4+ cells from becoming infected with human immunodeficiency virus.

15 This invention provides a method of treating a subject whose CD4+ cells are infected with human immunodeficiency virus which comprises administering to the subject an amount of one of the compounds described herein effective to inhibit binding of human immunodeficiency virus to CCR5 20 receptors on the surface of the subject's CD4+ cells so as to thereby treat the subject.

This invention provides a method of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor 25 which comprises:

- (a) immobilizing one of the compounds described herein on a solid support;
- (b) contacting the immobilized compound from step (a) with sufficient detectable CCR5 ligand to saturate all 30 binding sites for the CCR5 ligand on the immobilized

compound under conditions permitting binding of the CCR5 ligand to the immobilized compound so as to form a complex;

5 (c) removing any unbound CCR5 ligand;

10 (d) contacting the complex from step (b) with the agent; and

(e) detecting whether any CCR5 ligand is displaced from the complex, wherein displacement of detectable CCR5 ligand from the complex indicates that the agent binds to the compound so as to thereby identify the agent as one which inhibits binding of the CCR5 ligand to the CCR5 receptor.

This invention provides a method of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor which comprises:

15 (a) contacting one of the compounds described herein with sufficient detectable CCR5 ligand to saturate all binding sites for the CCR5 ligand on the compound under conditions permitting binding of the CCR5 ligand to the compound so as to form a complex;

20 (b) removing any unbound CCR5 ligand;

(c) measuring the amount of CCR5 ligand which is bound to the compound in the complex;

25 (d) contacting the complex from step (a) with the agent so as to displace CCR5 ligand from the complex;

(e) measuring the amount of CCR5 ligand which is bound to the compound in the presence of the agent; and

(f) comparing the amount of CCR5 ligand bound to the compound in step (e) with the amount measured in step

(c), wherein a reduced amount measured in step (e) indicates that the agent binds to the compound so as to thereby identify the agent as one which inhibits binding of the CCR5 ligand to the CCR5 receptor.

5

This invention also provides a method of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor which comprises:

- (a) immobilizing one of the compounds described herein on 10 a solid support;
- (b) contacting the immobilized compound from step (a) with the agent and sufficient detectable CCR5 ligand to saturate all binding sites for the CCR5 ligand on the compound under conditions permitting binding of the 15 CCR5 ligand to the immobilized compound so as to form a complex;
- (c) removing any unbound CCR5 ligand;
- (d) measuring the amount of detectable CCR5 ligand which is bound to the immobilized compound in the complex;
- 20 (e) measuring the amount of detectable CCR5 ligand which binds to the immobilized compound in the absence of the agent;
- (f) comparing the amount of CCR5 ligand which is bound to the immobilized compound in step (e) with the amount 25 measured in step (d), wherein a reduced amount measured in step (d) indicates that the agent binds to the compound so as to thereby identify the agent as one which inhibits binding of the CCR5 ligand to the CCR5 receptor.

30



This invention also provides a method of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor which comprises:

- (a) contacting one of the compounds described herein with the agent and sufficient detectable CCR5 ligand to saturate all binding sites for the CCR5 ligand on the compound under conditions permitting binding of the CCR5 ligand to the compound so as to form a complex;
- (b) removing any unbound CCR5 ligand;
- (c) measuring the amount of detectable CCR5 ligand which is bound to the compound in the complex;
- (d) measuring the amount of detectable CCR5 ligand which binds to the compound in the absence of the agent;
- (e) comparing the amount of CCR5 ligand which is bound to the compound in step (c) with the amount measured in step (d), wherein a reduced amount measured in step (c) indicates that the agent binds to the compound so as to thereby identify the agent as one which inhibits binding of the CCR5 ligand to the CCR5 receptor.

20

This invention provides a method of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor which comprises:

- a) immobilizing one of the compounds described herein on a solid support;
- b) contacting the immobilized compound from step a) with the agent dissolved or suspended in a known vehicle and measuring the binding signal generated by such contact;
- 30 c) contacting the immobilized compound from step a)

with the known vehicle in the absence of the compound and measuring the binding signal generated by such contact;

5           d) comparing the binding signal measured in step b) with the binding signal measured in step c), wherein an increased amount measured in step b) indicates that the agent binds to the compound so as to thereby identify the agent as one which binds to the CCR5 receptor.

10

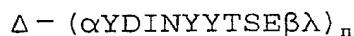
This invention provides a method of obtaining a composition which comprises:

15           (a) identifying a compound which inhibits binding of a CCR5 ligand to a CCR5 receptor according to one of the above methods; and

              (b) admixing the compound so identified or a homolog or derivative thereof with a carrier.

This invention provides a compound having the structure:

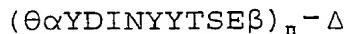
20



wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine; wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction;

wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ 5 ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction; wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group; wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined together by peptide bonds, further provided that at least 10 two tyrosines in the compound are sulfated, wherein  $\pi$  is an integer from 1 to 8,  $\Delta$  is a polymer, and the solid line represents up to 8 linkers which attach the structure in parentheses to  $\Delta$ .

15 This invention also provides a compound having the structure:



wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y 20 represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine; wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive 25 order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction; wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are 30 joined together by peptide bonds in consecutive order and

have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction; wherein  $\theta$  represents an amino group or an acetylated amino group;  
5 wherein all of  $\alpha$ , Y, D, I, N, Y, Y, T, S, E and  $\beta$  are joined together by peptide bonds, further provided that at least two tyrosines in the compound are sulfated, wherein  $\pi$  is an integer from 1 to 8,  $\Delta$  is a polymer, and the solid line represents up to 8 linkers which attach the structure in  
10 parentheses to  $\Delta$ .

This invention provides a compound having the structure:



wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine; wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids,  
15 they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction; wherein  $\beta$  represents from 0 to 333 amino acids, with the  
20 proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction; wherein  $\lambda$   
25 represents a carboxyl group or an amidated carboxyl group;  
30

wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined together by peptide bonds,

further provided that at least two tyrosines in the compound are sulfated, wherein  $n$  is an integer from 1 to 8,  $\Delta$  is a polymer, and the solid line represents up to 8 linkers which attach the structure in parentheses to  $\Delta$ .

This invention also provides a compound having the structure:

10  $(\Theta\alpha Y D I N Y Y T S E \beta)_n - \Delta$

wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine; wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction; wherein  $\beta$  represents from 0 to 333 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ 20 ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction; wherein  $\Theta$  represents an amino group or an acetylated amino group; wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined together by peptide bonds, further provided that at least 25 two tyrosines in the compound are sulfated, wherein  $n$  is an 30

integer from 1 to 8,  $\Delta$  is a polymer, and the solid line represents up to 8 linkers which attach the structure in parentheses to  $\Delta$ .

Brief Description of the Figures

Fig. 1 Effect of peptides on gp120<sub>JR-FL</sub> binding to CCR5.

5 L1.2-CCR5<sup>+</sup> cells were incubated with the biotinylated gp120<sub>JR-FL</sub>/CD4-IgG2 complex in the presence of different concentration of peptides (a) S-3/10/14, S-10/14, S-10, S-14 or (b) P-3/10/14, P-10/14, SR-2/12, SR-10/14, TS-10/14. The extent of complex binding in the absence of peptide was defined as 100% (m.f.i. ~40±5). Binding in the presence of peptide is expressed as a percentage of control. When CCR5-negative cells were used, binding of the gp120<sub>JR-FL</sub>/CD4-IgG2 complex was negligible (~10%, m.f.i. ~2±1). The 10 values shown are from a representative 15 experiment.

Fig. 2 Binding of the gp120/sCD4 complex to sulfated and phosphorylated peptides.

20 Biotinylated peptides were immobilized on a sensor chip and their ability to associate with gp120/sCD4 was analyzed by BIACore. RU values as a function of time were measured in the absence of peptide (gray dotted lines), in the presence of phosphorylated peptide (black dotted lines) or 25 in the presence of sulfated peptide (solid black lines). We performed binding analyses with the following proteins: (a) gp120<sub>JR-FL</sub>/sCD4, (b) gp120<sub>JR-FL</sub>, (c) sCD4, (d) DV3gp120<sub>JR-FL</sub>/sCD4, (e) gp120<sub>DH123</sub>/sCD4, (f) gp120<sub>DH123</sub>, (g) gp120<sub>LA1</sub>/sCD4 and 30

(h) gp120<sub>MAb</sub>

Fig. 3 Effect of peptides on MAb binding to CCR5.

L1.2-CCR5<sup>+</sup> cells were incubated with the anti-CCR5 MAbs in the presence of peptides. The extent of MAb binding in the absence of peptide was defined as 100% (m.f.i. ~50-400, depending on the MAb). Binding in the presence of peptide is expressed as a percentage of control. When CCR5-negative cells were used, binding of MAbs was negligible (m.f.i. ~2±1). Each data point represents the mean ± s.d. of three replicates.

Fig. 4 Binding of MAbs to sulfated and phosphorylated peptides.

Biotinylated peptides were immobilized on a sensor chip and their ability to associate with anti-CCR5 MAbs was analyzed by BIAcore. RU values as a function of time were measured in the absence of peptide (gray dotted lines), in the presence of phosphorylated peptide (black dotted lines) or in the presence of sulfated peptide (solid black lines). We performed binding analyses with (a) PA8, (b) PA10 and (c) 2D7.

Fig. 5 Effect of peptides on viral entry.

HeLa-CD4<sup>+</sup>CCR5<sup>+</sup> cells were infected with NLluc<sup>+</sup>env-pseudotyped with different viral envelopes in the presence of peptides. Luciferase activity (r.l.u.) was measured 48 h post-infection. The

extent of entry in the absence of peptide was defined as 100% (r.l.u.. ~25,000± 9,000). Background r.l.u. values were ~7±2. Each data point represents the mean ± s.d. of three replicates.

5 Fig. 6 CCR5 Nt peptide sequences and labels

10 The primary sequence of each peptide is indicated in the left column and the corresponding label is indicated in the right column. Sulfated tyrosine residues are designated by black boxes and white boxes designate phosphorylated tyrosine residues.

15 Fig. 7 Gp120/CD4 complex binding to CCR5 Nt sulfopeptides

20 Peptide 2-18 was bound to streptavidin-coated biosensor chips and gp120<sub>JR-FL</sub>/sCD4 (dotted line) or gp120<sub>JR-FL</sub>/CD4-IgG<sub>2</sub> (solid line) were flowed over the sensor chip surface. Resonance units (RU) were measured as a function of time using a Biacore X and reflect complex-peptide binding (a). Sulfopeptide 2-18 (solid symbols) or phosphopeptide 2-18(P) (clear symbols) were immobilized on streptavidin-coated ELISA plates and incubated with gp120/CD4-IgG<sub>2</sub> complexes. Gp120 25 proteins were derived from the R5 isolate JR-FL (squares), the R5X4 isolate DH123 (circles) and the X4 isolate LAI (diamonds). Complexes-peptide binding was detected by an HRP-conjugated goat anti-human IgG antibody. O.D. at 450 nm was

-19-

measured after addition of HRP substrate and is expressed as a function of CD4-IgG<sub>2</sub> concentration (b). Biotinylated sulfopeptide 2-18 was immobilized on streptavidin-coated plates and incubated with gp120/CD4-IgG<sub>2</sub> complex in the presence of increasing concentrations of: PA 8 (solid squares), TAK-779 (triangles), Rantes (inverted triangles), MIP-1 (diamonds), MIP-1 (circles) or SDF-1 (clear squares). Binding of the complexes to the peptide was detected by incubation with HRP-conjugated goat anti-human IgG antibody. O.D. at 450 nm was measured after addition of HRP substrate and percentage of binding was expressed as a function of inhibitor concentration.

Fig. 8: Binding of anti-CCR5 MAbs to CCR5 Nt peptides. Sulfopeptides (a) or phosphopeptides (b) were immobilized on streptavidin-coated ELISA plates and incubated with anti-CCR5 MAbs PA8 (solid squares), PA10 (clear circles), PA11 (solid circles), PA12 (solid diamonds) or PA14 (solid triangles). Binding of the MAbs to the peptides was detected by an HRP-conjugated goat anti-mouse IgG antibody. O.D. at 450 nm was measured after addition of HRP substrate and expressed as a function of MAb concentration.

Fig. 9: Binding of gp120<sub>JR-FL</sub>/CD4-IgG<sub>2</sub> to different CCR5 Nt-based peptides.

Streptavidin plates were coated with 2-18 (black squares), 10-18 (black circles), 8-15 (black diamonds), 6-16 (black stars), 10-15 (white square), 10-18(11A/18A) (black triangles). Plates  
5 were then incubated with gp120<sub>JR-FL</sub>/CD4-IgG<sub>2</sub> complex. Binding of the complex to the peptide was detected by an HRP-conjugated goat anti-human IgG antibody. O.D. at 450 nm was measured after addition of HRP substrate and expressed as a  
10 function of CD4-IgG<sub>2</sub> concentration (nM).

Fig. 10: Inhibition of gp120/CD4-IgG<sub>2</sub> complex binding to sulfo-peptides by anti-gp120 MAbs

Biotinylated sulfopeptide 2-18 was bound to streptavidin-coated biosensor chips and solutions of either gp120<sub>JR-FL</sub>/CD4-IgG<sub>2</sub> complex (black bars) or gp120<sub>DH123</sub>/CD4-IgG<sub>2</sub> complex (white bars) were flowed over the surface of the chip in the presence of different anti-gp120 MAbs. The names of the MAbs and the location of their epitopes are indicated along the x-axis. Resonance units (RU) were measured as a function of time using a Biacore X and reflect complex-peptide binding in the presence of the MAbs. Gp120/CD4-IgG<sub>2</sub> binding was calculated using the formula: (RU in the presence of MAbs) / (RU in the absence of MAbs) x100%. The values shown are from a sample experiment.  
20  
25

30 Fig. 11: Binding of gp120 mutants to sulfo-peptide and

wild type CCR5.

Sulfo-peptide 2-18 was immobilized on streptavidin-coated plates and incubated with a mixture of gp120-containing supernatants and CD4-IgG<sub>2</sub>. Peptide-complex binding was detected by an HRP-conjugated goat anti-human IgG antibody. O.D. at 450 nm was measured after addition of HRP substrate and normalized for binding of the gp120 mutants to CD4-IgG<sub>2</sub>. The dotted line represents the normalized value for the binding of the wild type gp120 to the peptide. The mutated amino acids and their locations in gp120 are indicated along the x-axis (a). L12-CCR5<sup>+</sup> cells were incubated with a mixture of gp120-containing supernatants and CD4-IgG<sub>2</sub>. Binding of the complex was detected by FACS analysis after addition of streptavidin-PE. Percentage of gp120/CD4-IgG<sub>2</sub> binding to CCR5 was normalized for gp120 binding to CD4-IgG<sub>2</sub>. The dotted line represents the normalized value for the binding of wild-type gp120 to the L12-CCR5<sup>+</sup> cells. The mutated amino acids and their locations in gp120 are indicated along the x-axis (b).

25

Fig. 12: Amino acid sequences of CCR5 Nt-based peptides. The peptides are named according to the positions of their first and last residues in the full-length sequence of CCR5. They contain either sulfotyrosines (black boxes) or phosphotyrosines

(white boxes) in positions 10 and 14. Residues Asp-11 and Glu-18 in peptide 10-18(11A/18A) are substituted for alanines. All peptides carry a carboxy terminal GAG spacer followed by a 5 biotinylated lysine.

Fig. 13: Amino acid conservation among R5 isolates. Envelope sequences from 25 R5 strains described in the HIV Database and retrieved from the 10 National Center for Biotechnology Information GenBank were aligned and percentage of conservation for the indicated residues was calculated and combined with results from Hung et al., 1999 (REF). Alanine mutants showing more 15 than 50 % decrease in sulfopeptide 2-18 binding compared to the wild type are highlighted in gray.

Detailed Description of the Invention

The plasmids CD4-IgG2-HC-pRcCMV and CD4-kLC-pRcCMV were deposited pursuant to, and in satisfaction of, the requirements of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms (the "Budapest Treaty") for the Purposes of Patent Procedure with the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, Virginia 20110-2209 under ATCC Accession Nos. 75193 and 75194, respectively.

10

The plasmids designated PPI4-tPA-gp120<sub>JRFL</sub> and PPI4-tPA-gp120<sub>LAI</sub> were deposited pursuant to, and in satisfaction of, the requirements of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure with the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, Virginia 20110-2209 under ATCC Accession Nos. 75431 and 75432, respectively. These plasmids were deposited with ATCC on March 12, 1993. These eukaryotic shuttle vectors contain the cytomegalovirus major immediate-early (CMV MIE) promoter/enhancer linked to the full-length HIV-1 envelope gene whose signal sequence was replaced with that derived from tissue plasminogen activator. In the vector, a stop codon has been placed at the gp120 C-terminus to prevent translation of gp41 sequences, which are present in the vector. The vector also contains an ampicillin resistance gene, an SV40 origin of replication and a DHFR gene whose transcription is driven by the  $\beta$ -globin promoter.

30

The monoclonal antibodies PA8, PA10, PA11, PA12, and PA14 were deposited pursuant to and in satisfaction of, the requirements of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the 5 Purposes of Patent Procedure with the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, Virginia 20110-2209 on December 2, 1998 under the following Accession Nos.: ATCC Accession No. HB-12605 (PA8), ATCC Accession No. HB-12607 (PA10), ATCC Accession No. HB-12608 10 (PA11), ATCC Accession No. HB-12609 (PA12), and ATCC Accession No. HB-12610 (PA14).

As used herein, the following standard abbreviations are used throughout the specification to indicate specific 15 amino acids:

|                                   |                     |
|-----------------------------------|---------------------|
| A=ala=alanine                     | R=arg=arginine      |
| N=asn=asparagine                  | D=asp=aspartic acid |
| C=cys=cysteine                    | Q=gln=glutamine     |
| 20 E=glu=glutamic acid            | G=gly=glycine       |
| H=his=histidine                   | I=ile=isoleucine    |
| L=leu=leucine                     | K=lys=lysine        |
| M=met=methionine                  | F=phe=phenylalanine |
| P=pro=proline                     | S=ser=serine        |
| 25 T=thr=threonine                | W=trp=tryptophan    |
| Y=tyr=tyrosine                    | V=val=valine        |
| B=asx=asparagine or aspartic acid |                     |
| Z=glx=glutamine or glutamic acid  |                     |

30 As used herein, the following standard abbreviations are

used throughout the specification to indicate specific nucleotides: C=cytosine; A=adenosine; T=thymidine; G= guanosine; and U=uracil.

5 This invention provides a compound comprising the structure:



wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y 10 represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine; wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order and 15 have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction; wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by 20 peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction;

wherein  $\theta$  represents an amino group or an acetylated amino 25 group; wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group; wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined together by peptide bonds; further provided that at least two tyrosines in the compound are sulfated.

30 In one embodiment of the above compound, the compound is

peptide which comprises consecutive amino acids having the sequence YDINYYTSE.

5 In one embodiment of the above compound, the tyrosines at positions 1 and 5 of the sequence YDINYYTSE are sulfated.

10 In one embodiment of the above compound,  $\alpha$  represents less than 9 amino acids. In another embodiment of the above compound,  $\alpha$  represents less than 8 amino acids. In another embodiment of the above compound,  $\alpha$  represents less than 7 amino acids. In another embodiment of the above compound,  $\alpha$  represents less than 6 amino acids. In another embodiment of the above compound,  $\alpha$  represents less than 5 amino acids. In another embodiment of the above compound,  $\alpha$  represents less than 4 amino acids. In another embodiment of the above compound,  $\alpha$  represents less than 3 amino acids. In another embodiment of the above compound,  $\alpha$  represents less than 2 amino acids. In another embodiment of the above compound,  $\alpha$  represents less than 1 amino acid.

15

20

In one embodiment of the above compound,  $\beta$  represents less than 17 amino acids. In one embodiment of the above compound,  $\beta$  represents less than 16 amino acids. In one embodiment of the above compound,  $\beta$  represents less than 15 amino acids. In one embodiment of the above compound,  $\beta$  represents less than 14 amino acids. In one embodiment of the above compound,  $\beta$  represents less than 13 amino acids. In one embodiment of the above compound,  $\beta$  represents less than 12 amino acids. In one embodiment of the above compound,  $\beta$  represents less than 11 amino acids. In one

25

30

embodiment of the above compound,  $\beta$  represents less than 10 amino acids. In one embodiment of the above compound,  $\beta$  represents less than 9 amino acids. In one embodiment of the above compound,  $\beta$  represents less than 8 amino acids.

5 In one embodiment of the above compound,  $\beta$  represents less than 7 amino acids. In one embodiment of the above compound,  $\beta$  represents less than 6 amino acids. In one embodiment of the above compound,  $\beta$  represents less than 5 amino acids. In one embodiment of the above compound,  $\beta$  represents less than 4 amino acids. In one embodiment of the above compound,  $\beta$  represents less than 3 amino acids. In one embodiment of the above compound,  $\beta$  represents less than 2 amino acids. In one embodiment of the above compound,  $\beta$  represents less than 1 amino acid.

15

This invention also provides a compound comprising the structure:

$\Theta\alpha YDINYYTSE\beta\lambda$

wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine; wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, 20 they are joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction; wherein  $\beta$  represents from 0 to 333 amino acids, with the proviso that 25 if there are more than 2 amino acids, they are joined by 30

peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction;

5 wherein θ represents an amino group or an acetylated amino group; wherein λ represents a carboxyl group or an amidated carboxyl group; wherein all of α,Y,D,I,N,Y,Y,T,S,E and β are joined together by peptide bonds; further provided that at least two tyrosines in the compound are sulfated.

10

In the compounds described herein and as exemplified above, the β in each compound may alternatively represent from 0 to 333 amino acids.

15 In one embodiment of the compounds described herein, β represents less than 300 amino acids. In another embodiment of the above compound, β represents less than 250 amino acids. In another embodiment of the above compound, β represents less than 200 amino acids. In another embodiment 20 of the above compound, β represents less than 150 amino acids. In another embodiment of the above compound, β represents less than 100 amino acids. In another embodiment of the above compound, β represents less than 75 amino acids. In another embodiment of the above compound, β represents less than 50 amino acids. In another embodiment 25 of the above compound, β represents less than 40 amino acids. In another embodiment of the above compound, β represents less than 35 amino acids. In another embodiment of the above compound, β represents less than 30 amino acids. In another embodiment of the above compound, β

represents less than 25 amino acids. In another embodiment of the above compound,  $\beta$  represents less than 20 amino acids. In another embodiment of the above compound,  $\beta$  represents less than 19 amino acids. In another embodiment 5 of the above compound,  $\beta$  represents less than 18 amino acids. In another embodiment of the above compound,  $\beta$  represents less than 17 amino acids. In another embodiment of the above compound,  $\beta$  represents less than 16 amino acids. In another embodiment of the above compound,  $\beta$  10 represents less than 15 amino acids. In another embodiment of the above compound,  $\beta$  represents less than 14 amino acids. In another embodiment of the above compound,  $\beta$  represents less than 13 amino acids. In another embodiment of the above compound,  $\beta$  represents less than 12 amino 15 acids. In another embodiment of the above compound,  $\beta$  represents less than 11 amino acids.

In one embodiment of the above compound,  $\alpha$  represents less than 9 amino acids. In another embodiment of the above 20 compound,  $\alpha$  represents less than 8 amino acids. In another embodiment of the above compound,  $\alpha$  represents less than 7 amino acids. In another embodiment of the above compound,  $\alpha$  represents less than 6 amino acids. In another embodiment of the above compound,  $\alpha$  represents less than 5 amino 25 acids. In another embodiment of the above compound,  $\alpha$  represents less than 4 amino acids. In another embodiment of the above compound,  $\alpha$  represents less than 3 amino acids. In another embodiment of the above compound,  $\alpha$  represents less than 2 amino acids. In another embodiment 30 of the above compound,  $\alpha$  represents less than 1 amino acid.

The CCR5 amino acid sequence is the following and is set forth in SEQ ID NO:1:

|    |   |     |
|----|---|-----|
| 1  | MDYQVSSPIYDINYTSEPCQKINVQIAARLLPPLYSLV        |     |
| 41 | FIFGFVGNMLVILILINCKRLKSMTDIYLLNL AISDLFFL     |     |
| 5  | 81 LTVPFWAHYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFII   |     |
|    | 121 LLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFAS  |     |
|    | 161 LPGIIFTRSQKEGLHYTCSSHFPYSQYQFWKNFQTLKIVI  |     |
|    | 201 LGLVLPLLVMVICYSGILKTLR CRNEKKRRAVRLIFTI   |     |
|    | 241 MIVYFLFWAPYNIVLLLNTFQEFG LNNCSSNRLDQAMQ   |     |
| 10 | 281 VTETLGMTHCCINPIIYAFVGEKFRNYLLVFFQKHI AKRF |     |
|    | 321 CKCCSIFQQEAPERASSVYTRSTGEQEISVGL          | 352 |

The CCR5 nucleotide sequence is the following and is set forth in SEQ ID NO:2:

|    |   |  |
|----|---|--|
| 15 | 1 GAATTCCCC AACAGAGCCA AGCTCTCCAT CTAGTGGACA GGGAAAGCTAG CAGCAAACCT     |  |
|    | 61 TCCCTTCACT ACAAAACCTTC ATTGCTTGGC CAAAAAGAGA GTTAATTCAA TGTAGACATC   |  |
|    | 121 TATGTAGGCA ATTAAAAACC TATTGATGTA TAAAACAGTT TGCATTCACTG GAGGGCAACT  |  |
|    | 181 AAATACATTC TAGGACTTTA TAAAAGATCA CTTTTTATTT ATGCACAGGG TGGAACAAGA   |  |
|    | 241 TGGATTATCA AGTGTCAAGT CCAATCTATG ACATCAATTA TTATACATCG GAGCCCTGCC   |  |
| 20 | 301 AAAAAATCAA TGTGAAGCAA ATCGCAGCCC GCCTCCTGCC TCCGCTCTAC TCACTGGTGT   |  |
|    | 361 TCATCTTGG TTTTGTGGGC AACATGCTGG TCATCCTCAT CCTGATAAAC TGCAAAAGGC    |  |
|    | 421 TGAAGAGCAT GACTGACATC TACCTGCTCA ACCTGGCCAT CTCTGACCTG TTTTCCTTC    |  |
|    | 481 TTACTGTCCC CTTCTGGCT CACTATGCTG CCGCCCAGTG GGACTTTGGA AATACAATGT    |  |
|    | 541 GTCAACTCTT GACAGGGCTC TATTTTATAG GCTTCCTCTC TGGAATCTTC TTCATCATCC   |  |
| 25 | 601 TCCTGACAAT CGATAGGTAC CTGGCTGTG TCCATGCTGT GTTGCTTTA AAAGCCAGGA     |  |
|    | 661 CGGTCACCTT TGGGGTGGTG ACAAGTGTGA TCACTTGGGT GGTGGCTGTG TTTGCGTCTC   |  |
|    | 721 TCCCAGGAAT CATCTT TACC AGATCTCAA AAGAAGGTCT TCATTACACC TGCAGCTCTC   |  |
|    | 781 ATTTTCCATA CAGTCAGTAT CAATTCTGGA AGAATTCCA GACATTAAAG ATAGTCATCT    |  |
|    | 841 TGGGGCTGGT CCTGCCGCTG CTTGTCATGG TCATCTGCTA CTCGGGAATC CTAAAAACTC   |  |
| 30 | 901 TGCTTCGGTG TCGAAATGAG AAGAAGAGGC ACAGGGCTGT GAGGCTTATC TTCACCATCA   |  |
|    | 961 TGATTGTTA TTTTCTCTTC TGGGCTCCCT ACAACATTGT CCTTCTCCTG AACACCTTCC    |  |
|    | 1021 AGGAATTCTT TGGCCTGAAT AATTGCAGTA GCTCTAACAG GTTGGACCAA GCTATGCAGG  |  |
|    | 1081 TGACAGAGAC TCTTGGGATG ACGCACTGCT GCATCAACCC CATCATCTAT GCCTTTGTGCG |  |
|    | 1141 GGGAGAAGTT CAGAAACTAC CTCTTAGTCT TCTTCCAAAA GCACATTGCC AAACGCTTCT  |  |

1201 GCAAATGCTG TTCTATTTTC CAGCAAGAGG CTCCCGAGCG AGCAAGCTCA GTTTACACCC  
1261 GATCCACTGG GGAGCAGGAA ATATCTGTGG GCTTGTGACA CGGACTCAAG TGGGCTGGTG  
1321 ACCCAGTCAG AGTTGTGCAC ATGGCTTAGT TTTCATACAC AGCCTGGGCT GGGGGT

5 The YDINYYTSE sequence corresponds to amino acid residues 10-18 of the CCR5 sequence set forth above.

As used herein, "CCR5" is a chemokine receptor which binds members of the CC group of chemokines and whose amino acid 10 sequence comprises that provided in Genbank Accession Number 1705896 and related polymorphic variants. The nucleotide sequence comprises that provided in Genbank Accession Number X91492. In one embodiment, the above compound may correspond to the extracellular portion of 15 CCR5. The first 31 amino acids of CCR5 correspond to the extracellular portion of CCR5. Accordingly, the extracellular portion extends from the methionine at position number 1 to the arginine at position number 31 of SEQ ID NO:1. In another embodiment, the above compound may 20 correspond to the amino-terminal portion of CCR5. As used herein, "N-terminus" or amino-terminus means the sequence of amino acids spanning the initiating methionine and the first transmembrane region.

25 As used herein, "H<sub>2</sub>N" refers to the N-terminus or amino-terminus. As used herein, "COOH" refers to the C-terminus or carboxy-terminus.

30 Various tyrosines of the compounds described herein may be sulfated. These include but are not limited to the tyrosines at positions 3, 10 and 14 of amino acid sequence

set forth in SEQ ID NO:1. Accordingly, in one embodiment, the tyrosines at positions 10 and 14 are sulfated. In another embodiment, the tyrosines at positions 3 and 14 are sulfated. In another embodiment, the tyrosines at positions 5 3 and 10 are sulfated. In another embodiment, the tyrosines at positions 3, 10 and 14 are sulfated. Other tyrosines in the sequence set forth in SEQ ID NO:1 may also be sulfated.

10 This invention provides a composition comprising one of the compounds described herein and a detectable marker attached thereto. In one embodiment of the composition, the detectable marker is biotin. In one embodiment of the composition, the detectable marker is attached at the C-terminus of the compound.

15

The compounds of the subject invention may also be isolated or purified. In one embodiment the compound is labeled with a detectable marker. As used herein, chemical "labels" include radioactive isotopes, fluorescent groups and 20 affinity moieties such as biotin that facilitate detection of the labeled peptide. Other chemical labels are well-known to those skilled in the art. Methods for attaching chemical labels to peptides are well-known to the skilled artisan.

25

As used herein, "peptide" and "polypeptide" are used to denote two or more amino acids linked by a peptidic bond between the  $\alpha$ -carboxyl group of one amino acid and the  $\alpha$ -amino group of the next amino acid. Peptides may be 30 produced by solid-phase synthetic methods that are well-

known to those skilled in the art. In addition to the above set of twenty amino acids that are used for protein synthesis in vivo, peptides may contain additional amino acids, including but not limited to hydroxyproline, 5 sarcosine, and  $\gamma$ -carboxyglutamate. The peptides may contain modifying groups including but not limited to sulfate and phosphate moieties. Peptides can be comprised of L- or D- amino acids, which are mirror-image forms with differing optical properties. Peptides containing D-amino acids have 10 the advantage of being less susceptible to proteolysis in vivo.

Peptides may be synthesized in monomeric linear form, cyclized form or as oligomers such as branched multiple 15 antigen peptide (MAP) dendrimers (Tam et al. Biopolymers 51:311, 1999). Nonlinear peptides may have increased binding affinity by virtue of their restricted conformations and/or oligomeric nature. Peptides may also be produced using recombinant methods as either isolated 20 peptides or as a portion of a larger fusion protein that contains additional amino acid sequences.

Peptides may be chemically conjugated to proteins by a variety of well-known methods. Such peptide-protein 25 conjugates can be formulated with a suitable adjuvant and administered parenterally for the purposes of generating polyclonal and monoclonal antibodies to the peptides of interest. Alternatively, unconjugated peptides can be formulated with adjuvant and administered to laboratory 30 animals for the purposes of generating antibodies. Methods

for generating and isolating such antibodies are well-known to those skilled in the art.

This invention provides derivatives of the above compound.

5 As used herein, a "derivative" peptide is one whose amino acid sequence is nonidentical to the reference peptide but which possesses functionally similar binding properties. Derivative peptides may also contain N-terminal, C-terminal and/or internal insertions, deletions, or substitutions of

10 amino acids, with the proviso that such insertions, deletions and substitutions do not abrogate the binding properties of the peptide. Derivative peptides include peptides modified with chemical labels to facilitate detection. Derivative peptides include branched and

15 cyclized peptides.

As used herein, "sulfopeptides" are peptides that contain sulfate moieties attached to one or more amino acids, such as tyrosine. In "sulfo-tyrosines", a sulfate group replaces

20 the *para*-hydroxyl group located on tyrosine side-chain.

As used herein, "phosphopeptides" are peptides that contain phosphate moieties attached to one or more amino acids, such a tyrosine. In "phospho-tyrosines", a phosphate group replaces the *para*-hydroxyl group located on tyrosine side-chain.

The peptides of the subject invention may be sulfated when synthesized or they may be subsequently sulfated. For

30 example, means of sulfating the peptides include chemical

sulfation or enzymatic sulfation. One skilled in the art would know how to employ these and other techniques to sulfate the compound.

- 5 This invention provides a composition which comprises a carrier and an amount of one of the compounds described herein effective to inhibit binding of HIV-1 to a CCR5 receptor on the surface of a CD4+ cell.
- 10 The carriers include but are not limited to an aerosol, intravenous, oral or topical carrier. Accordingly. The invention provides the above composition adapted for aerosol, intravenous, oral or topical application.
- 15 This invention provides the above compositions and a pharmaceutically acceptable carrier. Pharmaceutically acceptable carriers are well known to those skilled in the art. Such pharmaceutically acceptable carriers may include but are not limited to aqueous or non-aqueous solutions,
- 20 suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or
- 25 suspensions, saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers such as
- 30 those based on Ringer's dextrose, and the like.

Preservatives and other additives may also be present, such as, for example, antimicrobials, antioxidants, chelating agents, inert gases and the like.

5 As used herein, "composition" means a mixture. The compositions include but are not limited to those suitable for oral, rectal, intravaginal, topical, nasal, ophthalmic, or parenteral administration to a subject. As used herein, "parenteral" includes but is not limited to subcutaneous, 10 intravenous, intramuscular, or intrasternal injections or infusion techniques.

As used herein, "administering" may be effected or performed using any of the methods known to one skilled in the art. 15 The methods may comprise intravenous, intramuscular or subcutaneous means. As used herein, "effective dose" means an amount in sufficient quantities to either treat the subject or prevent the subject from becoming infected with HIV-1. A person of ordinary skill in the art can perform 20 simple titration experiments to determine what amount is required to treat the subject.

This invention provides a method of inhibiting human immunodeficiency virus infection of a CD4+ cell which also 25 carries a CCR5 receptor on its surface which comprises contacting the CD4+ cell with an amount of one of the compounds described herein effective to inhibit binding of human immunodeficiency virus to the CCR5 receptor so as to thereby inhibit human immunodeficiency virus infection of 30 the CD4+ cell. As used herein, "inhibits" means that the

amount is reduced. In a preferred embodiment, inhibits means that the amount is reduced 100%.

5 In one embodiment of this method, the CD4+ cell is present in a subject and the contacting is effected by administering the compound to the subject.

10 This invention provides a method of preventing CD4+ cells of a subject from becoming infected with human immunodeficiency virus which comprises administering to the subject an amount of one of the compounds described herein effective to inhibit binding of human immunodeficiency virus to CCR5 receptors on the surface of the CD4+ cells so as to thereby prevent the subject's CD4+ cells from 15 becoming infected with human immunodeficiency virus.

20 This invention provides a method of treating a subject whose CD4+ cells are infected with human immunodeficiency virus which comprises administering to the subject an amount of one of the compounds described herein effective to inhibit binding of human immunodeficiency virus to CCR5 receptors on the surface of the subject's CD4+ cells so as to thereby treat the subject.

25 As used herein, human immunodeficiency virus includes but is not limited to HIV-1, which is the human immunodeficiency virus type-1. HIV-1 includes but is not limited to extracellular virus particles and the forms of HIV-1 found in HIV-1 infected cells.

As used herein, "HIV-1 infection" means the introduction of HIV-1 genetic information into a target cell, such as by fusion of the target cell membrane with HIV-1 or an HIV-1 envelope glycoprotein<sup>+</sup> cell. The target cell may be a bodily 5 cell of a subject. In the preferred embodiment, the target cell is a bodily cell from a human subject.

As used herein, "inhibiting HIV-1 infection" means the reduction of the amount of HIV-1 genetic information 10 introduced into a target cell population as compared to the amount that would be introduced without the composition.

In the above methods, the compound may be administered by various routes including but not limited to aerosol, 15 intravenous, oral or topical route. The administration may comprise intralesional, intraperitoneal, intramuscular or intravenous injection; infusion; liposome-mediated delivery; topical, intrathecal, gingival pocket, per rectum, intrabronchial, nasal, oral, ocular or otic 20 delivery. In a further embodiment, the administration includes intrabronchial administration, anal, intrathecal administration or transdermal delivery. In another embodiment, the compound is administered hourly, daily, weekly, monthly or annually. In another embodiment, the 25 effective amount of the compound comprises from about 0.000001 mg/kg body weight to about 100 mg/kg body weight.

The administration may be constant for a certain period of time or periodic and at specific intervals. The compound 30 may be delivered hourly, daily, weekly, monthly, yearly

(e.g. in a time release form) or as a one time delivery. The delivery may be continuous delivery for a period of time, e.g. intravenous delivery.

5 The carrier may be a diluent, an aerosol, a topical carrier, an aqueous solution, a nonaqueous solution or a solid carrier.

10 The effective amount of the compound may comprise from about 0.000001 mg/kg body weight to about 100 mg/kg body weight. In one embodiment, the effective amount may comprise from about 0.001 mg/kg body weight to about 50 mg/kg body weight. In another embodiment, the effective amount may range from about 0.01 mg/kg body weight to about 15 10 mg/kg body weight. The actual effective amount will be based upon the size of the compound, the biodegradability of the compound, the bioactivity of the compound and the bioavailability of the compound. If the compound does not degrade quickly, is bioavailable and highly active, a 20 smaller amount will be required to be effective. The effective amount will be known to one of skill in the art; it will also be dependent upon the form of the compound, the size of the compound and the bioactivity of the compound. One of skill in the art could routinely perform 25 empirical activity tests for a compound to determine the bioactivity in bioassays and thus determine the effective amount.

30 The compound of the present invention may be delivered locally via a capsule which allows sustained release of the

agent or the peptide over a period of time. Controlled or sustained release compositions include formulation in lipophilic depots (e.g., fatty acids, waxes, oils). Also comprehended by the invention are particulate compositions 5 coated with polymers (e.g., poloxamers or poloxamines) and the agent coupled to antibodies directed against tissue-specific receptors, ligands or antigens or coupled to ligands of tissue-specific receptors. Other embodiments of the compositions of the invention incorporate particulate 10 forms protective coatings, protease inhibitors or permeation enhancers for various routes of administration, including parenteral, pulmonary, nasal and oral.

In one embodiment of the above methods, the subject is 15 infected with HIV-1 prior to administering the compound to the subject. In one embodiment of the above methods, the subject is not infected with HIV-1 prior to administering the compound to the subject. In one embodiment of the above methods, the subject is not infected with, but has been 20 exposed to, human immunodeficiency virus.

In one embodiment of the above methods, the effective amount of the compound comprises from about 1.0 ng/kg to about 100 mg/kg body weight of the subject. In another 25 embodiment of the above methods, the effective amount of the compound comprises from about 100 ng/kg to about 50 mg/kg body weight of the subject. In another embodiment of the above methods, the effective amount of the compound comprises from about 1  $\mu$ g/kg to about 10 mg/kg body weight 30 of the subject. In another embodiment of the above methods,

the effective amount of the compound comprises from about 100  $\mu\text{g}/\text{kg}$  to about 1  $\text{mg}/\text{kg}$  body weight of the subject.

5 The dose of the composition of the invention will vary depending on the subject and upon the particular route of administration used. Dosages can range from 0.1 to 100,000  $\mu\text{g}/\text{kg}$ . Based upon the composition, the dose can be delivered continuously, such as by continuous pump, or at periodic intervals. For example, on one or more separate 10 occasions. Desired time intervals of multiple doses of a particular composition can be determined without undue experimentation by one skilled in the art.

15 As used herein, "effective dose" means an amount in sufficient quantities to either treat the subject or prevent the subject from becoming infected with HIV-1. A person of ordinary skill in the art can perform simple titration experiments to determine what amount is required to treat the subject.

20 In one embodiment of the above method, the subject is a human being. As used herein, "subject" means any animal or artificially modified animal capable of becoming HIV-infected. Artificially modified animals include, but are 25 not limited to, SCID mice with human immune systems. The subjects include but are not limited to mice, rats, dogs, guinea pigs, ferrets, rabbits, and primates. In the preferred embodiment, the subject is a human being.

30 This invention provides a vaccine which comprises the

compound described herein. Vaccines comprising the sulfopeptides and a suitable adjuvant could be administered to a subject for the purposes of generating antibodies or other immune responses that are of therapeutic or prophylactic value. For example, the vaccines could be administered for the purpose of generating in the subject antibodies that bind CCR5 and inhibit its ability to mediate HIV entry and infection, thereby protecting the subject from HIV infection or disease progression. The vaccines may also comprise a suitable adjuvant. The vaccine may also comprises a suitable carrier.

The subject invention has various applications which includes HIV treatment such as treating a subject who has become afflicted with HIV. As used herein, "afflicted with HIV-1" means that the subject has at least one cell which has been infected by HIV-1. As used herein, "treating" means either slowing, stopping or reversing the progression of an HIV-1 disorder. In the preferred embodiment, "treating" means reversing the progression to the point of eliminating the disorder. As used herein, "treating" also means the reduction of the number of viral infections, reduction of the number of infectious viral particles, reduction of the number of virally infected cells, or the amelioration of symptoms associated with HIV-1. Another application of the subject invention is to prevent a subject from contracting HIV. As used herein, "contracting HIV-1" means becoming infected with HIV-1, whose genetic information replicates in and/or incorporates into the host cells. Another application of the subject invention is to

treat a subject who has become infected with HIV-1. As used herein, "HIV-1 infection" means the introduction of HIV-1 genetic information into a target cell, such as by fusion of the target cell membrane with HIV-1 or an HIV-1 envelope glycoprotein<sup>+</sup> cell. The target cell may be a bodily cell of a subject. In the preferred embodiment, the target cell is a bodily cell from a human subject. Another application of the subject invention is to inhibit HIV-1 infection. As used herein, "inhibiting HIV-1 infection" means reducing the amount of HIV-1 genetic information introduced into a target cell population as compared to the amount that would be introduced without said composition.

This invention provides a method of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor which comprises:

- (a) immobilizing one of the compounds described herein on a solid support;
- (b) contacting the immobilized compound from step (a) with sufficient detectable CCR5 ligand to saturate all binding sites for the CCR5 ligand on the immobilized compound under conditions permitting binding of the CCR5 ligand to the immobilized compound so as to form a complex;
- (c) removing any unbound CCR5 ligand;
- (d) contacting the complex from step (b) with the agent; and
- (e) detecting whether any CCR5 ligand is displaced from the complex, wherein displacement of detectable CCR5 ligand from the complex indicates that the agent binds

to the compound so as to thereby identify the agent as one which inhibits binding of the CCR5 ligand to the CCR5 receptor.

5 This invention provides a method of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor which comprises:

- (a) contacting one of the compounds described herein with sufficient detectable CCR5 ligand to saturate all 10 binding sites for the CCR5 ligand on the compound under conditions permitting binding of the CCR5 ligand to the compound so as to form a complex;
- (b) removing any unbound CCR5 ligand;
- (c) measuring the amount of CCR5 ligand which is bound to 15 the compound in the complex;
- (d) contacting the complex from step (a) with the agent so as to displace CCR5 ligand from the complex;
- (e) measuring the amount of CCR5 ligand which is bound to the compound in the presence of the agent; and
- 20 (f) comparing the amount of CCR5 ligand bound to the compound in step (e) with the amount measured in step (c), wherein a reduced amount measured in step (e) indicates that the agent binds to the compound so as to thereby identify the agent as one which inhibits 25 binding of the CCR5 ligand to the CCR5 receptor.

This invention also provides a method of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor which comprises:

- 30 (a) immobilizing one of the compounds described herein on

a solid support;

(b) contacting the immobilized compound from step (a) with the agent and detectable CCR5 ligand under conditions permitting binding of the CCR5 ligand to the immobilized compound so as to form a complex;

5 (c) removing any unbound CCR5 ligand;

(d) measuring the amount of detectable CCR5 ligand which is bound to the immobilized compound in the complex;

(e) measuring the amount of detectable CCR5 ligand which binds to the immobilized compound in the absence of the agent;

10 (f) comparing the amount of CCR5 ligand which is bound to the immobilized compound in step (e) with the amount measured in step (d), wherein a reduced amount measured in step (d) indicates that the agent binds to the compound so as to thereby identify the agent as one which inhibits binding of the CCR5 ligand to the CCR5 receptor.

15

20 In one embodiment of the above method, the amount of the detectable CCR5 ligand in step (a) and step (e) is sufficient to saturate all binding sites for the CCR5 ligand on the compound.

25 This invention also provides a method of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor which comprises:

(a) contacting one of the compounds described herein with the agent and detectable CCR5 ligand under conditions permitting binding of the CCR5 ligand to the compound

30

so as to form a complex;

- (b) removing any unbound CCR5 ligand;
- (c) measuring the amount of detectable CCR5 ligand which is bound to the compound in the complex;
- 5 (d) measuring the amount of detectable CCR5 ligand which binds to the compound in the absence of the agent;
- (e) comparing the amount of CCR5 ligand which is bound to the compound in step (c) with the amount measured in step (d), wherein a reduced amount measured in step 10 (c) indicates that the agent binds to the compound so as to thereby identify the agent as one which inhibits binding of the CCR5 ligand to the CCR5 receptor.

In one embodiment of the above method, the amount of the 15 detectable CCR5 ligand in step (a) and step (d) is sufficient to saturate all binding sites for the CCR5 ligand on the compound.

In one embodiment of the above method the solid support is 20 a microtiter plate well. In another embodiment, the solid support is a bead. In a further embodiment, the solid support is a surface plasmon resonance sensor chip. The surface plasmon resonance sensor chip can have pre-immobilized streptavidin. In one embodiment, the surface 25 plasmon resonance sensor chip is a BIACore™ chip.

In one embodiment of the above methods, the detectable CCR5 ligand is labeled with a detectable marker. In another embodiment of the above methods, the CCR5 ligand is 30 detected by contacting it with another compound which is

both capable of detecting the CCR5 ligand and is detectable. The detectable markers include those described above.

5 This invention provides a method of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor which comprises:

- a) immobilizing one of the compounds described herein on a solid support;
- 10 b) contacting the immobilized compound from step a) with the agent dissolved or suspended in a known vehicle and measuring the binding signal generated by such contact;
- c) contacting the immobilized compound from step a) with the known vehicle in the absence of the compound and measuring the binding signal generated by such contact;
- 15 d) comparing the binding signal measured in step b) with the binding signal measured in step c), wherein an increased amount measured in step b) indicates that the agent binds to the compound so as to thereby identify the agent as one which binds to the CCR5 receptor.

25 In one embodiment of the above method, the solid support is a surface plasmon resonance sensor chip. In another embodiment of the above method, the binding signal is measured by surface plasmon resonance.

30 This invention provides a method of obtaining a composition

which comprises:

- (a) identifying a compound which inhibits binding of a CCR5 ligand to a CCR5 receptor according to one of the above methods; and
- 5 (b) admixing the compound so identified or a homolog or derivative thereof with a carrier.

The invention provides agents identified in the screen. Such agents may have utility in treating HIV-1 infection or 10 other CCR5-mediated diseases, which include rheumatoid arthritis, asthma, multiple sclerosis, psoriasis, atherosclerosis and other inflammatory diseases.

In one embodiment of the above methods, the CCR5 ligand is 15 a complex comprising an HIV-1 envelope glycoprotein and a CD4-based protein. The HIV-1 envelope glycoproteins include but are not limited to gp120, gp140 or gp160. The CD4-based proteins include but are not limited to soluble CD4 or CD4-IgG2.

20

As used herein, "CD4" means the mature, native, membrane-bound CD4 protein comprising a cytoplasmic domain, a hydrophobic transmembrane domain, and an extracellular domain that binds to the HIV-1 gp120 envelope glycoprotein.

25

As used herein, "HIV-1 envelope glycoprotein" means the HIV-1 encoded protein which comprises the gp120 surface protein, the gp41 transmembrane protein and oligomers and precursors thereof. As used herein, "CD4-based protein" means any protein comprising at least one sequence of amino 30 acid residues corresponding to that portion of CD4 which is

required for CD4 to form a complex with the HIV-1 gp120 envelope glycoprotein. As used herein, "CD4-IgG2" means a heterotetrameric CD4-human IgG2 fusion protein encoded by the expression vectors deposited under ATCC Accession 5 Numbers 75193 and 75194.

In one embodiment of the above methods, the CCR5 ligand is a chemokine. The chemokines include but are not limited to RANTES, MIP-1 $\alpha$  or MIP-1 $\beta$ . As used herein, "RANTES", "MIP-1 $\alpha$ ", 10 and "MIP-1 $\beta$ " denote members of the chemokine superfamily of proteins that direct the activation and migration of leukocytes and other cells involved in the inflammation. RANTES, MIP-1 $\alpha$  and MIP-1 $\beta$  are known to bind CCR5 and induce 15 signaling. Their peptide sequences have been described (Wells et al. J. Leukocyte Biology, 59:53-60, 1996).

In one embodiment of the above methods, the CCR5 ligand is an antibody. In one embodiment, the antibody is PA8 (ATCC Accession No. HB-12605). In another embodiment, the 20 antibody is PA10 (ATCC Accession No. 12607). In another embodiment, the antibody is PA11 (ATCC Accession No. HB-12608). In another embodiment, the antibody is PA12 (ATCC Accession No. HB-12609).

25 This invention provides a compound having the structure:

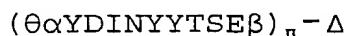


wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, 30 each I represents an isoleucine; and each N represents an

asparagine; wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction; wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction; wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group; wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined together by peptide bonds, further provided that at least two tyrosines in the compound are sulfated, wherein  $\pi$  is an integer from 1 to 8,  $\Delta$  is a polymer, and the solid line represents up to 8 linkers which attach the structure in parentheses to  $\Delta$ .

20

This invention also provides a compound having the structure:



wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine; wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive

order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction; wherein  $\beta$  represents from 0 to 13 amino acids, with the 5 proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction; wherein  $\theta$  10 represents an amino group or an acetylated amino group; wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined together by peptide bonds, further provided that at least two tyrosines in the compound are sulfated, wherein  $\pi$  is an integer from 1 to 8, 15  $\Delta$  is a polymer, and the solid line represents up to 8 linkers which attach the structure in parentheses to  $\Delta$ .

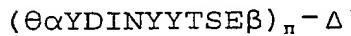
This invention provides a compound having the structure:



20 wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine; wherein  $\alpha$  represents from 0 to 9 amino acids, 25 with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction; 30 wherein  $\beta$  represents from 0 to 333 amino acids, with the

proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending 5 therefrom in the carboxy terminal direction; wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group; wherein all of  $\alpha$ , Y, D, I, N, Y, Y, T, S, E and  $\beta$  are joined together by peptide bonds, further provided that at least 10 two tyrosines in the compound are sulfated, wherein  $\pi$  is an integer from 1 to 8,  $\Delta$  is a polymer, and the solid line represents up to 8 linkers which attach the structure in parentheses to  $\Delta$ .

15 This invention also provides a compound having the structure:



wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, 20 each I represents an isoleucine; and each N represents an asparagine; wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ 25 ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction; wherein  $\beta$  represents from 0 to 333 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and 30 have a sequence identical to the sequence set forth in SEQ

ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction; wherein  $\theta$  represents an amino group or an acetylated amino group; wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined 5 together by peptide bonds, further provided that at least two tyrosines in the compound are sulfated, wherein  $n$  is an integer from 1 to 8,  $\Delta$  is a polymer, and the solid line represents up to 8 linkers which attach the structure in parentheses to  $\Delta$ .

10

The polymer of the above compounds includes but is not limited to the following: a linear lysine polymer; a branched lysine polymers; a linear arginine polymer; a branched arginine polymer; and polyethylene glycol (PEG), a 15 linear acetylated lysine polymer, a branched acetylated lysine polymer, a linear chloroacetylated lysine polymer and a branched chloroacetylated lysine polymer.

The above compounds can be produced by various methods 20 known to those skilled in the art, including but not limited to the following. Methods for producing synthetic multimeric peptides such as multiple antigen peptides, synthetic polymeric constructs, and branched lysine oligopeptides are well known to those skilled in the art 25 (Spetzler and Tam, Int. J. Pept. Prot. Res. 45:78, 1995; Yai et al., J. Virol., 69:320, 1995; Okuda et al., J. Mol. Recognit. 6:101, 1993). For example, radially branched peptides can be produced by performing standard solid-phase 30 peptide synthesis methods using branched lysine skeletons on 4-(oxy-methyl)-phenylactamidomethyl or other suitable

solid resin. Peptide chains are elongated in parallel in a stepwise fashion using optimized t-butyloxycarbonyl/benzyl chemistry as described (Sabatier et al., Biochemistry 32:2763, 1993). Peptides are liberated from the resin, 5 purified by reversed-phase chromatography over a C18 or other suitable column and characterized by analytical HPLC and mass spectroscopy. In another approach, monomeric peptides are synthesized, purified, and then covalently coupled to lysine copolymers using N-succinimidyl maleimido 10 carboxylate chemistry. In another approach, the peptides can also be made in the form of affinity type multimers. For example, peptides may be synthesized with an affinity tag such as biotin. These affinity tagged peptides can then be mixed with affinity ligands capable of binding 15 multimerically, such as streptavidin. Other site-specific ligation chemistries are known to the skilled artisan.

This invention provides a compound comprising the structure:

20  $\Theta\alpha Y D \Pi \Pi Y \Pi \Pi E \beta \lambda$

wherein each E represents a glutamic acid, each D represents an aspartic acid, and each Y represents a tyrosine;

25 wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction;

30 wherein  $\beta$  represents from 0 to 13 amino acids, with the

-55-

proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending 5 therefrom in the carboxy terminal direction; wherein Θ represents an amino group or an acetylated amino group; wherein λ represents a carboxyl group or an amidated carboxyl group; wherein Π represents any amino acid, 10 wherein all of α, Y, D, Π, Π, Y, Π, Π, Π, E and β are joined together by peptide bonds; further provided that at least two tyrosines in the compound are sulfated.

15 In one embodiment of this compound, the compound comprises amino acids in addition to those in the YDΠΠYΠΠΠE peptide, and such amino acids correspond to those present in the CCR5 receptor sequence set forth in SEQ ID NO:1, yet an amino acid may be replaced with a homologous amino acid. 20 The sequence YDΠΠYΠΠΠE corresponds to amino acid residues 10-18 of the sequence set forth in SEQ ID NO:1. For example, if the peptide has one additional amino acid on its N terminal end, then the sequence could be IYDΠΠYΠΠΠE or alternatively, the I could be replaced with G, A, V or 25 L.

In one embodiment of the above compound, the compound is a peptide which comprises consecutive amino acids having the sequence YDΠΠYΠΠΠE.

In one embodiment of the above compound, the tyrosines at positions 1 and 5 of the sequence YDΠΠYΠΠIE are sulfated.

As used herein, "homologous amino acids" are those which 5 have chemically similar side chains. For example, aliphatic side chains (G, A, V, L and I); aromatic side chains (F, Y and W); basic aide chains (K, R and H); acidic side chains (D and E); amide side chains (N and Q); aliphatic hydroxyl-containing side chains (S and T); sulfur-containing side 10 chains (C and M). Homology between amino acids may also be drawn on other bases, such as size, polarity, hydrogen bonding potential, hydrophilicity and hydrophobicity. Proline differs from the above amino acids in that it contains a secondary rather than primary imino group. 15 Accordingly, proline may be considered an imino group. Substitution of proline with another amino acid (e.g. G, A or S) can increase the flexibility of a peptide. Conversely, substitution of another amino acid with a proline can stabilize a desired conformation.

20

This invention provides a compound comprising the structure:

$\Theta\alpha Y D I N Y Y T S E \beta \lambda$

wherein each T represents a threonine, each S represents a 25 serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine;

wherein  $\alpha$  represents from 0 to 9 amino acids, with the 30 proviso that if there are more than 2 amino acids, they are

joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction;

5 wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending

10 therefrom in the carboxy terminal direction;

wherein  $\theta$  represents an amino group or an acetylated amino group; wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group;

wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined

15 together by peptide bonds;

further provided that at least two tyrosines in the compound are sulfated,

wherein any amino acid except for the Y at position 1, D at

20 position 2, Y at position 5 and E at position 9 may be replaced with a homologous amino acid.

In one embodiment of the above compound, with respect to replacing homologous amino acids, any I amino acid residue may be replaced with a G,A,V or L amino acid residue. In

25 one embodiment of the above compound, any N amino acid residue may be replaced with a Q amino acid residue. In one embodiment of the above compound, any Y amino acid residue may be replaced with a F or W amino acid residue. In one embodiment of the above compound, any T amino acid residue

30 may be replaced with a S amino acid residue. In one

embodiment of the above compound, any S amino acid residue may be replaced with a T amino acid residue. In one embodiment of the above compound, any C may be replaced with M, S, T, A, G, N, or Q.

5

In one embodiment, a C amino acid residue within the  $\beta$  region of the compound may be replaced with any other amino acid.

10 This invention provides an agent which binds to an epitope of HIV-1 gp120, which epitope comprises amino acid residues R298, N301, T303, I322, D324, I325, R326, I420, K421, Q422, W427, thereby inhibiting binding of HIV-1 gp120 to a CCR5 chemokine receptor.

15

The above amino acid numbering is per HIV-1 strain HxB2. (Genbank Accession No. AAB50262). Amino acids D324, I325 and R326 are derived from HIV-1 strain JR-FL (Genbank Accession No. AAB05604).

20

25

30

The amino acid sequence (SEQ ID NO:17) for HIV-1 HxB2 gp120 is set forth below:

1 MRVKEKYQHL WRWGWRWGTM LLGMLMICS A TEKLWVTVYY GVPVWKEATT TLFCASDAKA  
61 YDTEVHNWVA THACVPTDPN P QEVVVLVNVT ENFNWKNDM VEQMHEIDIIS LWDQSLKPCV  
5 121 KLTPLCVSLK CTDLKNDTNT NSSSGRMIME KGEIKNCSFN ISTSIRGKVQ KEYAFFYKLD  
181 IIPIDNDTTS YKLTSCNTSV ITQACPKVSF EPIPIHYCAP AGFAILKCNN KTFNGTGPCT  
241 NVSTVQCTHG IRPVVSTQLL LNGSLAEEEV VIRSVNFTDN AKTIIIVQLNT SVEINCTRPN  
301 NNTRKRIRIQ RGPGRAFVTI GKIGNMRQAH CNISRAKWNN TLKQIASKLR EQFGNNKTII  
361 FKQSSGGDPE IVTHSFNCGG EFFYCNSTQL FNSTWFnSTW STEGSNNTEG SDTITLPCR  
10 421 KQIINMWQKV GKAMYAPPIS GQIRCSSNIT GLLLTRDGGN SNNESEIFRP GGGDMRDNW  
481 SELYKYKVVK IEPLGVAPTK AKRRVVQREK R

The amino acid sequence (SEQ ID NO:16) for HIV-1 JR-FL gp120 is set forth below:

1 MRVKGIRKSY QYLWKGGTLL LGILMICS A EKLWVTVYYG VPVWKEATT TLFCASDAKAY  
61 DTEVHNWAT HACVPTDPNP QEVVLENVTE HFNMWKNNMV EQMQEDIISL WDQSLKPCVK  
121 LTPLCVTLNC KDVNATNTTN DSEGTMERGE IKNCSFNITT SIRDEVQKEY ALFYKLDVVP  
181 IDNNNTSYRL ISCDTSVITQ ACPKISFEP I PIHYCAPAGF AILKCNDKTF NGKGPCKNVS  
20 241 TVQCTHGIRP VVSTQLLLNG SLAEEEVVIR SDNFTNNAKT IIVQLKESVE INCTRPNNT  
301 RKSIIHGPGF AFYTTGEIIG DIRQAHCNIS RAKWNDTLKQ IVIKLREQFE NKTIVFNHSS  
361 GGDPEIVMHS FNCGGEFFY C NSTQLFNSTW NNNTEGSNN EGNTITLPCR IKQIINMWQE  
421 VGKAMYAPPI RGQIRCSSNI T GLLLTRDGG INENGTEIFR PGGGDMRDNW RSELYKYKVV  
481 KIEPLGVAPT KAKRRVVQRE KR

25

This invention provides the above agent, wherein the epitope is altered or masked by an alanine substitution of at least one of the amino acid residues R298, N301, T303, I322, D324, I325, R326, I420, K421, Q422 and W427.

30

This invention provides an agent which binds to an epitope of HIV-1 gp120, which epitope comprises amino acid residues R298, N301, T303, I322, D324, I325, R326, I420, K421, Q422,

W427, thereby inhibiting HIV-1 infection of a CD4+ CCR5+ cell.

This invention provides the above agent, wherein the 5 epitope is altered or masked by an alanine substitution of at least one of the amino acid residues R298, N301, T303, I322, D324, I325, R326, I420, K421, Q422 and W427.

In one embodiment of any of the above agents, the agent is 10 a peptide. In one embodiment of any of the above agents, the peptide comprises consecutive amino acids having the sequence YDINYYTSE. In one embodiment at least two tyrosines in the compound are sulfated. In one embodiment, the tyrosines at positions 1 and 5 of the sequence 15 YDINYYTSE are sulfated.

In one embodiment of any of the above agents, the agent is an antibody or portion of an antibody. In one embodiment of any of the above agents, the agent is a nonpeptidyl agent. 20 In one embodiment of any of the above agents, the agent is a peptidyl agent.

This invention provides a method of inhibiting HIV-1 infection of a CD4+CCR5+ cell which comprises contacting 25 the cell with an amount of an agent of the subject invention effective to bind to HIV-1 gp120, so as to thereby inhibit HIV-1 infection of the CD4+ CCR5+ cell.

This invention provides a compound having one of the 30 following structures:

$\Delta$ -( $\alpha$ YDINYYTSE $\beta$  $\lambda$ ), ( $\theta$  $\alpha$ YDINYYTSE $\beta$ )- $\Delta$ , or  $\Delta$ -( $\alpha$ YDINYYTSE $\beta$ )- $\Delta$

wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine;

5 wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and  
10 have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction;

15 wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction;

20 wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group;

wherein  $\theta$  represents an amino group or an acetylated amino group;

wherein all of  $\alpha$ , Y, D, I, N, Y, Y, T, S, E and  $\beta$  are joined together by peptide bonds,

25 further provided that at least two tyrosines in the compound are sulfated,

wherein  $\Delta$  is a molecule that self-oligomerizes, and the solid line represents a peptide linker or a peptide, disulfide, or other chemical bond.

As used herein "peptide linker" is a peptide comprising consecutive amino acids having a sequence which includes but is not limited to GAG, SGGRGG and QSTRGGASGGG or repeating units thereof. One skilled in the art would know 5 other flexible peptide linkers.

In one embodiment of the above compound, the peptide that self-oligomerizes contains alpha-helical regions capable of forming coiled coils.

10

The  $\alpha$ -helical coiled coil (48) is probably the most widespread subunit oligomerization motif found in proteins (48-52). It is a type of protein structure consisting of two to five amphipathic  $\alpha$ -helices that "coil" around each 15 other in a left-handed supertwist (48-52). The sequences of coiled are characterized by a heptad repeat of seven residues with a hydrophobic repeat of mostly apolar amino acids.

20

In one embodiment of the above compound, the peptide that self-oligomerizes is a peptide having a sequence of at least a portion HIV-1 gp41 heptad repeat sequence 1. In one embodiment, the HIV-1 gp41 heptad repeat sequence 1 is RQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARILVERYLKDQ (SEQ ID 25 NO:3).

In one embodiment of the above compound, the peptide that self-oligomerizes is a peptide having a sequence of at least a portion of an HIV-1 gp41 heptad repeat sequence 2.

30

In one embodiment, the HIV-1 gp41 heptad repeat sequence 2

is WMEWDREINNYTSЛИHSLIEESQNQQEKNEQELLE (SEQ ID NO:4).

In one embodiment of the above compound, the peptide that self-oligomerizes is a peptide having a sequence 5 corresponding to at least a portion of the leucine zipper region of transcription factor GCN4. In one embodiment, the sequence of the leucine zipper region of transcription factor GCN4 is HMKQLEDKVEELLSKNYHLENEVARLKKLVGER (SEQ ID NO:6).

10

In one embodiment of the above compound, the peptide that self-oligomerizes is a peptide having a sequence corresponding to at least a portion of the leucine zipper region of transcription factor GCN4. In one embodiment, the sequence is derived from the leucine zipper region of transcription factor GCN4. In one embodiment, the sequence forms trimeric coiled-coils. In one embodiment, the sequence is HMKQIEDKIEEILSKIYHIENEIARIKKLIGEV (SEQ ID NO:7).

20

In one embodiment of the above compound, the peptide that self-oligomerizes is a peptide having a sequence corresponding to at least a portion of a leucine zipper region of a human protein. The human protein includes but 25 is not limited to transcription activator c-fos, transcription activator c-jun, enzyme quiescent cell proline dipeptidase, macrophage scavenger receptor, salivary mucin (MUC7), or human quiescent cell proline dipeptidase (QPP). In one embodiment, the human protein is 30 QPP and the leucine zipper region has the sequence

LLTVEQALADFAELLRALRRDL (SEQ ID NO:5). In one embodiment, the transcription activator is c-fos and the leucine zipper region has the sequence LTDTLQAETDQLEDEKSALQTEIANLLKEKEKLEFILAAR (SEQ ID NO:8). In 5 one embodiment, the transcription activator is c-jun and the leucine zipper region has the sequence HMRRIARLEEKVCTLKAQNSELASTANMLREQVAQLKQKY (SEQ ID NO:9).

In one embodiment of the above compound, the peptide that 10 self-oligomerizes is a peptide having a sequence corresponding to that of at least a portion of an antibody. In one embodiment, the portion of the antibody comprises the heavy chain. In one embodiment, the portion of the antibody heavy chain comprises the heavy chain constant 15 region. In one embodiment, the portion of the antibody heavy chain comprises the hinge and Fc domains. In one embodiment, the portion of the antibody heavy chain comprises the Fc domain. In one embodiment, the portion of the antibody comprises the light chain. In one embodiment, the portion 20 comprises the light chain constant region. In one embodiment, the portion of the antibody comprises the heavy and light chains.

This invention provides a compound having one of the 25 following structures:

$\Delta$ -( $\alpha$ YDINYYTSE $\beta$  $\lambda$ ), ( $\theta$  $\alpha$ YDINYYTSE $\beta$ ) $-\Delta$ , or  $\Delta$ -( $\alpha$ YDINYYTSE $\beta$ ) $-\Delta$   
wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, 30 each I represents an isoleucine; and each N represents an

asparagine;

wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction;

wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction;

wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group;

wherein  $\theta$  represents an amino group or an acetylated amino group;

wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined together by peptide bonds,

further provided that at least two tyrosines in the compound are sulfated,

wherein  $\Delta$  is toxin, and the solid line represents a peptide linker or a peptide, disulfide, or other chemical bond.

In one embodiment of the above compound, the toxin is a radionuclide. In one embodiment, the radionuclide is an alpha-emitting isotope. The alpha-emitting isotope includes but is not limited to  $^{225}\text{Ac}$ ,  $^{211}\text{At}$ ,  $^{212}\text{Bi}$ , or  $^{213}\text{Bi}$ . In one embodiment, the radionuclide is a beta-emitting isotope.

The beta-emitting isotope includes but is not limited to

$^{186}\text{Rh}$ ,  $^{188}\text{Rh}$   $^{90}\text{Y}$ ,  $^{131}\text{I}$ , or  $^{67}\text{Cu}$ . In one embodiment, the radionuclide may be emitting Auger and low energy electron. The radionuclide includes but is not limited to  $^{131}\text{I}$ ,  $^{125}\text{I}$  or  $^{77}\text{Br}$ .

5

In one embodiment of the above compound, the toxin is a chemical toxin. The chemical toxin may be a peptidyl chemical toxin. The peptidyl chemical toxin includes but is not limited to ricin. The chemical toxin may be a 10 nonpeptidyl chemical toxin. The nonpeptidyl chemical toxin includes but is not limited to calicheamycin.

This invention provides a compound having one of the following structures:

15  $\Delta$ - $(\alpha\text{YDINYYTSE}\beta\lambda)$ ,  $(\theta\alpha\text{YDINYYTSE}\beta) - \Delta$ , or  $\Delta$ - $(\alpha\text{YDINYYTSE}\beta) - \Delta$ ,

wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an 20 asparagine;

wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ 25 ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction;

wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and 30 have a sequence identical to the sequence set forth in SEQ

ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction;  
wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group;

5 wherein  $\theta$  represents an amino group or an acetylated amino group;  
wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined together by peptide bonds,  
further provided that at least two tyrosines in the  
10 compound are sulfated,  
wherein  $\Delta$  is molecule with anti-HIV activity, and the solid line represents a peptide linker or a peptide, disulfide, or other chemical bond.

15 In one embodiment of the above compound, the molecule with anti-HIV activity is a CD4-immunoglobulin fusion protein. In one embodiment, the CD4-immunoglobulin fusion protein is CD4-IgG2, wherein the CD4-IgG2 comprises two heavy chains and two lights chains, wherein the heavy chains are encoded  
20 by an expression vector designated CD4-IgG2HC-pRcCMV (ATCC Accession No. 75193) and the light chains are encoded by an expression vector designated CD4-kLC-pRcCMV (ATCC Accession No. 75194).

25 In one embodiment of the above compound, the molecule with anti-HIV activity is a compound which retards gp41 from adopting a conformation capable of mediating fusion of HIV-1 to a CD4+ cell by binding noncovalently to an epitope on a gp41 fusion intermediate. In one embodiment, the compound  
30 comprises a peptide selected from the group consisting of

T-20 (SEQ ID NO: 10), DP107 (SEQ ID NO: 11), N34 (SEQ ID NO: 12), C28 (SEQ ID NO: 13), N34(L6)C28 (SEQ ID NO: 14), and T1249 (SEQ ID NO:15).

5 As used herein, "T-20" and "DP178" are used interchangeably to denote a peptide having the following amino acid sequence: YTSLIHSILIEESQNQQEKNEQELLELDKWASLWNWF (SEQ ID NO:10) and as described [53,54].

10 DP107 has the following amino acid sequence:  
NNLLRAIEAQHQHLLQLTVWGIKQLQARILAVERYLKDQ (SEQ ID NO:11)

N34 has the following amino acid sequence:  
SGIVQQQNLLRAIEAQHQHLLQLTVWGIKQLQAR (SEQ ID NO:12) .

15 C28 has the following amino acid sequence:  
WMEWDREINNYTSLIHSILIEESQNQQEK (SEQ ID NO:13)

N34(L6)C28 has the following amino acid sequence:  
20 SGIVQQQNLLRAIEAQHQHLLQLTVWGIKQLQARSGGRGGWMEWDREINNYTSLIHSILIEESQNQQEK (SEQ ID NO:14)

T1249 has the following amino acid sequence :  
WQEWEQKITALLEQAQIQQQEKNEYELQKLDKWASLWEWF (SEQ ID NO:15)

25 This invention provides the above compound wherein the molecule with anti-HIV activity is a CCR5 chemokine receptor targeting agent.

30 In one embodiment, the CCR5 chemokine receptor targeting

agent is an antibody or portion of an antibody. In one embodiment, the antibody includes but is not limited to PA8 (ATCC Accession No. HB-12605), PA10 (ATCC Accession No. 12607), PA11 (ATCC Accession No. HB-12608), PA12 (ATCC 5 Accession No. HB-12609), and PA14 (ATCC Accession No. HB-12610). In one embodiment, the antibody is PA14 (ATCC Accession No. HB-12610).

10 The antibody may be a monoclonal antibody or polyclonal antibody. The monoclonal antibody may be a human, humanized or chimeric antibody. This invention provides humanized forms of the above antibodies.

15 As used herein, "humanized" describes antibodies wherein some, most or all of the amino acids outside the CDR regions are replaced with corresponding amino acids derived from human immunoglobulin molecules. In one embodiment of the humanized forms of the antibodies, some, most or all of the amino acids outside the CDR regions have been replaced 20 with amino acids from human immunoglobulin molecules but where some, most or all amino acids within one or more CDR regions are unchanged. Small additions, deletions, insertions, substitutions or modifications of amino acids are permissible as long as they would not abrogate the 25 ability of the antibody to bind a given antigen. Suitable human immunoglobulin molecules would include IgG1, IgG2, IgG3, IgG4, IgA and IgM molecules. A "humanized" antibody would retain a similar antigenic specificity as the original antibody, i.e., in the present invention, the 30 ability to bind CCR5.

One skilled in the art would know how to make the humanized antibodies of the subject invention. Various publications, several of which are hereby incorporated by reference into this application, also describe how to make humanized 5 antibodies. For example, the methods described in United States Patent No. 4,816,567 (55) comprise the production of chimeric antibodies having a variable region of one antibody and a constant region of another antibody.

10 United States Patent No. 5,225,539 (56) describes another approach for the production of a humanized antibody. This patent describes the use of recombinant DNA technology to produce a humanized antibody wherein the CDRs of a variable region of one immunoglobulin are replaced with the CDRs 15 from an immunoglobulin with a different specificity such that the humanized antibody would recognize the desired target but would not be recognized in a significant way by the human subject's immune system. Specifically, site directed mutagenesis is used to graft the CDRs onto the 20 framework.

Other approaches for humanizing an antibody are described in United States Patent Nos. 5,585,089 (57) and 5,693,761 (58) and WO 90/07861 which describe methods for producing 25 humanized immunoglobulins. These have one or more CDRs and possible additional amino acids from a donor immunoglobulin and a framework region from an accepting human immunoglobulin. These patents describe a method to increase the affinity of an antibody for the desired antigen. Some 30 amino acids in the framework are chosen to be the same as

the amino acids at those positions in the donor rather than in the acceptor. Specifically, these patents describe the preparation of a humanized antibody that binds to a receptor by combining the CDRs of a mouse monoclonal 5 antibody with human immunoglobulin framework and constant regions. Human framework regions can be chosen to maximize homology with the mouse sequence. A computer model can be used to identify amino acids in the framework region which are likely to interact with the CDRs or the specific 10 antigen and then mouse amino acids can be used at these positions to create the humanized antibody.

The above patents 5,585,089 and 5,693,761, and WO 90/07861 (59) also propose four possible criteria which may be used in 15 designing the humanized antibodies. The first proposal was that for an acceptor, use a framework from a particular human immunoglobulin that is unusually homologous to the donor immunoglobulin to be humanized, or use a consensus framework from many human antibodies. The second proposal 20 was that if an amino acid in the framework of the human immunoglobulin is unusual and the donor amino acid at that position is typical for human sequences, then the donor amino acid rather than the acceptor may be selected. The third proposal was that in the positions immediately 25 adjacent to the 3 CDRs in the humanized immunoglobulin chain, the donor amino acid rather than the acceptor amino acid may be selected. The fourth proposal was to use the donor amino acid reside at the framework positions at which the amino acid is predicted to have a side chain atom 30 within 3Å of the CDRs in a three dimensional model of the

antibody and is predicted to be capable of interacting with the CDRs. The above methods are merely illustrative of some of the methods that one skilled in the art could employ to make humanized antibodies.

5

This invention provides the above compound, wherein the portion of the antibody is a Fab fragment of the antibody. This invention provides the above compound, wherein the portion of the antibody comprises the variable domain of 10 the antibody. This invention provides the above compound, wherein the portion of the antibody comprises a complementarity determining region or CDR portion of the antibody. The monoclonal antibody includes but is not limited to an IgG, IgM, IgD, IgA, or IgE monoclonal 15 antibody.

This invention provides the above compound, wherein the molecule with anti-HIV activity is a chemokine or chemokine derivative. The chemokine includes but is not limited to 20 RANTES, MIP-1 $\alpha$ , MIP-1 $\beta$ , SDF-1 or other chemokine which blocks HIV-1 infection. The chemokine derivative includes but is not limited to Met-RANTES, AOP-RANTES, RANTES 9-68, or NNY-RANTES.

25 The molecule may also be a non-chemokine agent capable of binding to chemokine receptors and inhibiting fusion of HIV-1 to CD4 $^{+}$  cells. The non-chemokine agents include, but are not limited to, chemokine fragments and chemokine derivatives and analogues. In one embodiment, the agent 30 does not include naturally occurring chemokines. The non-

chemokine agents include multimeric forms of the chemokine fragments and chemokine derivatives and analogues or fusion molecules which contain chemokine fragments, derivatives and analogues linked to other molecules. In one embodiment, the non-chemokine agents do not include bicyclams and their derivatives as described in U.S. Patent No. 5,021,409, issued June 4, 1991, the content of which is incorporated by reference into this application. Some bicyclam derivatives have been previously described with 10 antiviral activities (60, 61).

In an embodiment of this invention, the non-chemokine agent is an oligopeptide. In another embodiment, the non-chemokine agent is a polypeptide. In still another embodiment, the non-chemokine agent is an antibody or a portion thereof. Antibodies against the chemokine receptor may easily be generated by routine experiments. It is also within the level of ordinary skill to synthesize fragments of the antibody capable of binding to the chemokine receptor. In a further embodiment, the non-chemokine agent is a nonpeptidyl agent such as TAK-779 (64) or AMD3100 (65).

Non-chemokine agents which are purely peptidyl in composition can be either chemically synthesized by solid-phase methods (62) or produced using recombinant technology in either prokaryotic or eukaryotic systems. The synthetic and recombinant methods are well known in the art.

30 Non-chemokine agents which contain biotin or other

nonpeptidyl groups can be prepared by chemical modification of synthetic or recombinant chemokines or non-chemokine agents. One chemical modification method involves periodate oxidation of the 2-amino alcohol present on 5 chemokines or non-chemokine agents possessing serine or threonine as their N-terminal amino acid (63). The resulting aldehyde group can be used to link peptidyl or non-peptidyl groups to the oxidized chemokine or non-chemokine agent by reductive amination, hydrazine, or other 10 chemistries well known to those skilled in the art.

This invention provides a compound having one of the following structures:

$\Delta$ -( $\alpha$ YDINYYTSE $\beta$  $\lambda$ ), ( $\theta$  $\alpha$ YDINYYTSE $\beta$ ) $-\Delta$ , or  $\Delta$ -( $\alpha$ YDINYYTSE $\beta$ ) $-\Delta$   
15 wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine;  
20 wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending 25 therefrom in the amino terminal direction;  
wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ 30 ID NO: 1 beginning with the P at position 19 and extending

therefrom in the carboxy terminal direction;  
wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group;  
wherein  $\theta$  represents an amino group or an acetylated amino group;  
5 wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined together by peptide bonds,  
further provided that at least two tyrosines in the compound are sulfated,  
10 wherein  $\Delta$  is peptidyl or nonpeptidyl agent, and the solid line represents a peptide linker, or a peptide, disulfide, or other chemical bond.

15 In one embodiment, the  $\Delta$  is a nonpeptidyl agent, and the nonpeptidyl agent polyethylene glycol.

This invention will be better understood from the Experimental Details that follow. However, one skilled in the art will readily appreciate that the specific methods 20 and results discussed are merely illustrative of the invention as described more fully in the claims that follow thereafter.

EXPERIMENTAL DETAILS

First Series of Experiments

A. Materials

5 Purified recombinant CD4-IgG2 protein was produced by Progenics Pharmaceuticals, Inc. from plasmids CD4-IgG2-HC-pRcCMV and CD4-kLC-pRcCMV as described (Allaway et al. AIDS Res. Hum. Retroviruses 11:533, 1995). Soluble CD4 is commercially available (NEN Life Science Products, Boston, MA). Anti-CCR5 MAb 2D7 was purchased from Pharmingen (San-Diego, CA).

10 The plasmids designated PPI4-tPA-gp120<sub>JR-FL</sub>-V3<sup>(-)</sup> and PPI4-tPA-gp120<sub>DH123</sub> were prepared as described (Hasel et al, US Patents 5,869,624 and 5,886,163). Monomeric gp120 glycoproteins were produced in CHO cells stably transfected with the PPI4-tPA-gp120 plasmids and purified to homogeneity as described (Hasel et al. US Patents 5,869,624 and 5,886,163; Trkola et al. Nature 384:184, 1996). The 15 antibodies designated PA8, PA10, PA12 and PA14 were prepared by growing the corresponding hybridoma cell line in mouse ascites and isolating the antibody using protein A affinity chromatography as described (Olson et al. J.Virol. 73:4145, 1999). L1.2-CCR5<sup>+</sup> cells were cultured as described 20 (Olson et al. J.Virol. 73:4145, 1999).

25 Peptides containing different segments of the CCR5 Nt were custom-synthesized by solid-phase fluorenylmethoxycarbonyl chemistry using phospho- and sulfo-tyrosine precursors as building blocks where indicated (Figure 6). Biotinylated

versions of peptides S-10/14 and P-10/14 incorporated a C-terminal GAG spacer preceding a biotinylated lysine. Following cleavage from the resin, peptides were purified by reverse-phase chromatography on C18 columns (Vydac, 5 Hesperia, CA) and analyzed by HPLC and mass spectroscopy. Figure 6 describes the different peptides that were used in this study.

Binding of gp120 to CCR5

10 A gp120/CD4 complex formed from monomeric gp120 (100nM) and biotinylated CD4-IgG2 (50nM) was added to  $1 \times 10^6$  L1.2-CCR5<sup>+</sup> cells in the presence of different concentrations of peptide (Olson *et al.* J.Virol. 73:4145, 1999). CD4-IgG2 is tetrameric and therefore binds four molecules of gp120, 15 which increases binding of the complex to CCR5 (Allaway *et al.* AIDS Res. Hum. Retroviruses 11:533, 1995). The mean fluorescence intensity (m.f.i.) was measured by flow cytometry after addition of phycoerythrin (PE)-labeled streptavidin (Becton Dickinson, San Jose, CA). Inhibition of 20 gp120/CCR5 binding was calculated: (m.f.i. with peptide)/(m.f.i. without peptide)  $\times 100\%$ .

It was first tested whether tyrosine-sulfated peptides spanning amino acids 2-18 of the CCR5 Nt could inhibit 25 binding of the gp120<sub>JR-FL</sub>/CD4-IgG2 complex to CCR5<sup>+</sup> cells. The HIV-1<sub>JR-FL</sub> isolate exclusively uses CCR5 as a co-receptor (Dragic *et al.* Nature 381:667, 1996). Only peptides S-3/10/14 and S-10/14 inhibited complex binding to the cells in a dose-dependent manner (Fig. 1a). Peptides S-10 and S-30 14 had no inhibitory activity, even at the highest

concentrations (Fig. 1a). Peptide TS-10/14, spanning amino acids 10-14, did not inhibit gp120<sub>JR-FL</sub>/CD4-IgG2 binding to CCR5<sup>+</sup> cells, despite the presence of two sulfo-tyrosine residues (Fig. 1b).

5

Tyrosine-phosphorylated peptides P-10/14 and P-3/10/14 did not inhibit gp120<sub>JR-FL</sub>/CD4-IgG2 binding to CCR5<sup>+</sup> cells (Fig. 1b). As further specificity controls we synthesized peptides containing the first seventeen residues of the 10 CCR5 Nt in random order with sulfo-tyrosines in positions 10 and 14 (SS-10/14) or in positions 2 and 12 (SS-2/12). Neither one of these peptides reduced gp120<sub>JR-FL</sub>/CD4-IgG2 binding to CCR5<sup>+</sup> cells, even at the highest concentrations (Fig. 1b).

15

Surface plasmon resonance measurements (BIAcore)

Streptavidin-coated sensor chips (BIAcore AB, Sweden) were conditioned with five injections of regeneration solution (1M NaCl, 50mM NaOH) and equilibrated with HBS-EP buffer 20 (10mM HEPES, 150mM NaCl, 3M EDTA, 0.005% polysorbate 20) as recommended by the manufacturer. Biotinylated peptides were then immobilized on the chip by injection of peptide (100nM) in HBS-EP buffer, followed by an injection of regeneration solution and equilibration with HBS-EP buffer. 25 400 resonance units (RU) of peptide were bound to the sensor chip surface. Solutions of the following proteins (100nM) were passed over the sensor chip surface: gp120, sCD4, gp120/sCD4, PA8, PA10 and 2D7. Surface plasmon resonance was monitored and displayed in arbitrary 30 resonance units (RU) as a function of time. Following

injection of each solution the chip was regenerated and equilibrated as described above.

Biotinylated peptide was attached to the streptavidin-coated gold surface of a sensor chip and solutions containing different gp120/sCD4 complexes were flowed over the immobilized peptide. Adsorption of the complex due to complex/peptide binding was detected by an increase in surface plasmon resonance signal (RU), which reports changes in the effective refraction index very near the gold surface of the sensor chip (Schuck Ann. Rev. Biophys Biomol Struct 26:541, 1997). For proteins of similar size, such as the different gp120/sCD4 complexes, RU plateau values are directly proportional to the amount of protein bound to the peptide.

Specific association of the gp120<sub>JR-FL</sub>/sCD4 complex with the sulfo-tyrosine-containing peptide bS-10/14 was accompanied by a significant increase in RU (Fig. 2a). The signal plateau but not the shape of the sensograms varied with gp120<sub>JR-FL</sub>/sCD4 concentration indicating that the peptide/complex interaction was dose-dependent (data not shown). The sensogram obtained with bP-10/14 is similar to the one obtained in the absence of peptide, indicating a complete lack of association of the phosphorylated peptide with the protein complex (Fig. 2a). Neither gp120<sub>JR-FL</sub> nor sCD4 alone produced a significant increase in RU, indicating that they did not associate with the immobilized peptides. (Fig. 2b,c). The gp120-ΔV3<sub>JR-FL</sub>/sCD4 complex was also unable to associate with the peptides (Fig. 2d).

To further ascertain the specificity of the peptide/complex association we performed BIACore analyses using envelope glycoproteins from HIV-1<sub>DH123</sub>, an R5X4 isolate, and HIV-1<sub>LAI</sub>, an X4 isolate (5). gp120<sub>DH123</sub>/sCD4 associated specifically 5 with the sulfated peptide, although the plateau RU values were lower than those observed with gp120<sub>JR-FL</sub>/sCD4 (Fig. 2e). We did not detect any binding of gp120<sub>DH123</sub>/sCD4 to the phosphorylated peptide (Fig. 2e), nor did gp120<sub>DH123</sub> alone associate with the peptides (Fig. 2f). Finally, gp120<sub>LAI</sub> with 10 or without sCD4 was not able to associate with either one of the peptides (Fig. 2g,h).

These methods could be readily modified to screen for agents that bind CCR5 or that block its interaction with 15 antibodies, gp120 or other ligands. For example, direct binding of the agents could be analyzed as described above, where the agent is substituted for the anti-CCR5 antibody or gp120/sCD4 complex. In another embodiment, the agent could be mixed or pre-incubated with the anti-CCR5 antibody 20 (or gp120/sCD4 complex) prior to passing the mixture over biosensor chips as described above.

Binding of MAbs to CCR5

L1.2-CCR5 cells ( $1 \times 10^6$ ) were incubated with anti-CCR5 MAb 25 (50nM)  $\pm$  peptide (100 $\mu$ M). MAb binding was detected using a PE-labeled goat anti-mouse antibody (Caltag Laboratories, Burlingam, CA). The m.f.i. value was measured by flow cytometry as described (Olson *et al.* J. Virol. 73:4145, 1999). MAb binding was calculated as above.

We determined whether the CCR5 Nt peptides could inhibit binding of a panel of anti-CCR5 Mabs to CCR5<sup>+</sup> cells. PA8 binding was reduced significantly by all wild-type peptides containing amino acids 2-18, regardless of tyrosine 5 modification (Fig. 3). BIAcore analysis confirmed that PA8 similarly and specifically associated with both sulfated and phosphorylated peptides (Fig. 4). Binding of PA12 to CCR5 was not inhibited by any of the peptides (Fig. 3). PA10 binding to CCR5 was inhibited only by S-3/10/14 (Fig. 10 3). PA10 was also observed to associate with bS-10/14 and to a lesser extent with bP-10/14 in BIAcore analysis (Fig. 4), which may be more sensitive than the gp120/CCR5-binding assay. Binding of 2D7 to CCR5 was not inhibited by any of the peptides (Fig. 3). No significant interaction was 15 observed between any CCR5 Nt peptide and Mab 2D7 (Figs. 3 and 4), whose epitope resides within the second extracellular loop on CCR5.

Single cycle HIV-1 entry assay

20 Nluc<sup>+</sup>env<sup>-</sup> particles pseudotyped with envelope glycoproteins from MuLV, HTLV-1 and HIV-1 strains JR-FL, HxB<sub>2</sub>, DH123, Gun-1 were made as described (Dragic et al. J. Virol. 72:279, 1998). Target cells (HeLa-CD4<sup>+</sup>CCR5<sup>+</sup> or U87-CD4<sup>+</sup>CCR5<sup>+</sup>) were 25 incubated with virus-containing supernatant fractions (100ng/ml p24) ± peptide (100μM) for 4 h. then washed and resuspended in culture media. After 48 hours the cells were lysed and luciferase activity (relative light units, r.l.u.) was measured using a standard kit (Promega, Madison, WI) as described (Dragic et al. J. Virol. 72:279, 30 1998). Viral entry was calculated: (r.l.u. with

peptide) / (r.l.u without peptide) x100%.

The ability of different CCR5 Nt peptides to inhibit HIV-1 entry into CD4<sup>+</sup>CCR5<sup>+</sup>CXCR4<sup>+</sup> cells was tested using a 5 luciferase-based single round of entry assay (5). Only peptides S-10/14 and S-3/10/14 inhibited the entry of the R5 isolate HIV-1<sub>JR-FL</sub> by approximately 50% in HeLa-CD4<sup>+</sup>CCR5<sup>+</sup> and U87MG-CD4<sup>+</sup>CCR5<sup>+</sup> (Fig. 5 and data not shown). We were 10 unable to inhibit the entry of the R5X4 isolates HIV-1<sub>DH123</sub> and HIV-1<sub>Gun-1</sub>, or of the X4 isolate HIV-1<sub>HXB2</sub>. The entry of MuLV and HTLV pseudotypes was also unaffected by the peptides (Fig. 5).

15 Screening assays

1) HIV-1 gp120/CD4-IgG2

Streptavidin-coated 96-well microtiter plates (NEN Life 20 Science Products, Boston, MA) are blocked with 200  $\mu$ l/well of 5% bovine serum albumin (Sigma, St. Louis, MO) in PBS buffer and washed with assay buffer (0.5% Tween 20, 1 % fetal bovine serum, and 2% BSA in PBS buffer). The plates are then incubated 1 hour at ambient temperature with 100  $\mu$ l/well of biotinylated CCR5 N-terminal sulfopeptide at a 25 concentration of 500  $\mu$ M in assay buffer. Following a wash step, the plates are incubated for 1 hour at ambient temperature with an HIV-1<sub>JR-FL</sub> gp120/CD4-IgG2 complex in the presence or absence of inhibitory agent. The plates are again washed and incubated for 30 minutes with a horseradish peroxidase-labeled goat antibody to human IgG 30 (Kirkegaard & Perry Laboratories, Gaithersburg, MD)

followed by addition of the TMB (3,3',5,5'-tetramethylbenzidine) chromogenic substrate (Pierce). The reaction is stopped by addition of 100  $\mu$ l/well of 2N  $H_2SO_4$  prior to colorimetric detection at a wavelength of 450 nm.

5 Wells without biotinylated peptide serve as negative controls. The percent inhibition of binding is calculated as  $[1 - (OD_{\text{with inhibitor}} - OD_{\text{control well}}) / (OD_{\text{without inhibitor}} - OD_{\text{control well}})] \times 100$ , where OD represents the average optical density observed for the indicated wells.

10

2) Anti-CCR5 antibodies

Streptavidin-coated microtiter plates are blocked and incubated with CCR5 N-terminal peptide as described above. Following a wash step, the plates are incubated for one 15 hour at ambient temperature with the anti-CCR5 antibody PA10 in the presence or absence of inhibitory agent. The plates are again washed and incubated for 30 minutes with a horseradish peroxidase-labeled goat antibody to mouse IgG (Kirkegaard & Perry Laboratories, Gaithersburg, MD) 20 followed by addition of TMB substrate for colorimetric detection as described above. The percent inhibition mediated by the inhibitory agent is calculated as described above.

25

Discussion

Tyrosine-modified peptides spanning the region of the CCR5 Nt that contains residues important for viral entry were synthesized. (Dragic *et al.* J. Virol. 72:279, 1998; Rabut 30 *et al.* J. Virol. 72:3464, 1998; Farzan *et al.* J. Virol.

72:1160, 1998; Dorantz et al. J. Virol. 71:6305, 1997). Interactions between the Nt peptides and gp120/CD4 complexes were characterized. Peptides containing sulfo-tyrosines in positions 10 and 14 efficiently inhibited 5 binding of gp120<sub>JR-FL</sub>/CD4 to CCR5. Substitution of the sulfate groups for phosphates, which are also negatively charged at physiological pH, rendered the Nt peptides inactive. Inhibition of gp120/CCR5 binding was dependent, therefore, on the presence of sulfate moieties and was not 10 simply due to non-specific electrostatic interactions between the peptide and the gp120/CD4 complex or the peptide and the cell surface. Inhibition of gp120/CCR5 binding was also dependent on the primary structure surrounding the sulfo-tyrosines since peptides with random 15 sequences of CCR5 amino acids 2-18 had no inhibitory activity. Additional Nt amino acids in the region 2-18 were important for activity since a shortened peptide containing just amino acids 10-14 was unable to inhibit gp120/CD4 binding, despite the presence of two sulfo-tyrosines. It 20 would be straightforward to define the minimum number of amino acids needed for activity by systematically synthesizing sulfopeptides intermediate in length between peptide 2-18 and peptide 10-14. Similarly, sulfopeptides that incorporate a greater portion of the CCR5 Nt could be 25 easily synthesized and tested for activity using the methods described herein.

Qualitative BIACore analyses allowed the demonstration of a highly specific, CD4-dependent interaction between a 30 tyrosine-sulfated Nt peptide and gp120<sub>JR-FL</sub>. No binding of

the protein complex to a tyrosine-phosphorylated peptide was observed. Only gp120s derived from isolates that use CCR5 as a co-receptor associated with the sulfated peptide. gp120<sub>DH123</sub>/CD4 binding was weaker than gp120<sub>JR-FL</sub>/CD4 binding, 5 suggesting that envelope glycoproteins from R5X4 isolates have a lower apparent affinity for CCR5 than envelope glycoproteins from R5 isolates. gp120<sub>LAI</sub>, derived from an isolate that only uses CXCR4, did not bind to the sulfated peptide. A V3 loop-deleted gp120<sub>JR-FL</sub> did not associate with 10 the sulfated peptide, just as this protein was unable to bind to full length CCR5 on the cell surface (Trkola *et al.* *Nature* 384:184, 1996).

The binding of the Nt peptides to several anti-CCR5 MAbs, 15 all of which recognize conformational epitopes in CCR5 and inhibit gp120/CCR5 binding were also studied. PA12 and 2D7 did not bind to any of the peptides. Binding of PA8 to the peptides was independent of tyrosine-modification whereas PA10 associated more with the sulfo-tyrosine-containing 20 peptide than with the phospho-tyrosine-containing peptide. It seems, therefore, that sulfo-tyrosines and phospho-tyrosines are relatively interchangeable for the purpose of MAb binding but that gp120/CD4 binding has an absolute requirement for sulfo-tyrosines. Relatively subtle 25 differences in size and geometry of sulfate and phosphate groups might be relevant for binding of the CCR5 Nt with gp120, which must not only accept the negative charge, but also coordinate, probably by hydrogen bonds, the tyrosine sulfate oxygens. The kinetics of MAb binding to the CCR5 Nt 30 peptides exhibited large apparent on rates and slow

-86-

apparent off rates, which also differed from our observations of gp120/CD4 binding kinetics.

None of the Nt peptides inhibited MuLV, HTLV and HIV-1<sub>HXB2</sub> envelope-mediated viral entry, which is not mediated by CCR5. In contrast, peptides S-10/14 and S-3/10/14 specifically inhibited the entry of the HIV-1<sub>JR-FL</sub> R5 strain in two different cell lines. The inhibition of HIV-1 entry by tyrosine-sulfated peptides was partial (~50%) but nonetheless striking given the difficulty of blocking this process with short, linear peptides (Jameson et al. Science 240:1335, 1988; Chan and Kim Cell 93:681:1998; Doranz et al. J. Exp. Med. 186:1395, 1997; Heveker et al. Current Biology 8:369, 1998; Eckert et al. Cell 99:1, 1999).

15

20

25

30

References for First Series of Experiments:

1. E. A. Berger et al. (1999) Annu Rev Immunol 17: 657-700.
2. A. Trkola et al. (1996) Nature 384: 184-187.
- 5 3. L. Wu et al. (1996) Nature 384: 179-183.
4. Y.J. Zhang et al. (1999) J. Virol 73: 3443-3448.
5. T. Dragic et al. (1998) J. Virol. 72: 279-285.
6. G.E. Rabut et al. (1998) J. Virol. 72:3464-3468.
7. M. Farzan et al. (1998) J. Virol. 72:1160-1164.
- 10 8. B.J. Doranz et al. (1997) J. virol. 71: 6305-6314.
9. M. Farzan et al. (1999) Cell 96:667-676.
10. P.A. Baeuerle et al. (1987) J. Cell Biol 105:2655-2664.
11. J.W. Kehoe et al. (2000) Chemistry & Biology 7:R57-15 R61.
12. M. Baba et al. (1988) Proc Natl Acad Sci USA 85: 6132-6136.
13. D. Schols et al. (1990) Virology 175: 556-561.
14. G. Roderiquez (1995) J. Virol 69: 2233-2239.
- 20 15. S.S. Hwang et al. (1991) Science 253: 71-74.
16. F. Safaiyan et al. (1999) J. Biol Chem 274: 36267-36273.
17. P.A. Baeuerle et al. (1986) Biochem Biophys Res Commun 141: 870-877.
- 25 18. W.C. Olson et al. (1999) J Virol 73:4145-4155.
19. G.P. Allaway et al. (1995) AIDS Res Hum Retroviruses 11: 533-539.
20. T. Dragic et al. (1996) Nature 381:667-673.
21. P. Schuck et al. (1997) Annu Rev Biophys Biomol Struct 30 26:541-566.

22. S. Lederman et al. (1989) *J. Immunol* 143:1149-1154.
23. S. Ohlson et al. (2000) *Trends in Biotechnology* 18: 49-52.
24. B.A. Jameson et al. (1988) *Science* 240: 1335-1339.
- 5 25. D.C. Chan et al. (1998) *Cell* 93: 681-684.
26. B.J. Doranz et al. (1997) *J. Exp Med* 186: 1395-1400.
27. N. Heveker et al. (1998) *Current Biology* 8:369-376.
28. D.M. Eckert et al. (1999) *Cell* 99: 1-20.
29. H. Sakaida et al. (1998) *J. Virol* 72: 9763-9770.
- 10 30. S.E. Kuhman et al. (1997) *J. Virol* 71: 8642-8656.
31. T.M. Ross et al. (1998) *J. Virol* 72: 1918-1924.
32. R.E. Atchison et al. (1996) *Science* 274: 1924-1926.
33. P.D. Bieniasz et al. (1997) *EMBO* 16: 2599-2609.
34. A.L. Edinger et al. (1999) *J. Virol* 73:4062-4073.
- 15 35. J. Rucker et al. (1996) *Cell* 87: 437-446.
36. L. Picard et al. (1997) *J. Virol* 71: 5003-5011.
37. T. Dragic et al. (2000) *Proc Natl Acad Sci* 10: 1073.
38. P.D. Kwong et al. (1998) *Nature* 393: 648-659.
39. C.D. Rizzuto et al. (1998) *Science* 280: 1949-1953.
- 20 40. P.D. Kwong et al. (2000) *J. Virol* 74: 1961-1972.
41. M. Moulard et al. (2000) *J. Virol* 74: 1948-1960.
42. A. Trkola et al. (1996) *Nature* 384: 184-186.
43. B. Labrosse et al. (1998) *J. Virol* 72: 6381-6388.
44. Spetzler et al. (1995) *J. Pept. Prot Res* 45: 78.
- 25 45. Yai et al. Et al. (1995) *J. Virol* 69: 320.
46. Okuda et al. (1993) *J. Mol. Recognit.* 6:101.
47. Sabatier et al. (1993) *Biochemistry* 32: 2763.
48. Frank et al (2000) *Journal of Biological Chemistry* 275. No. 16: 11672-11677.
- 30 49. C. Cohen et al (1990) *Proteins* 7:1-15.

50. A. Lupas (1996) *Trends Biochem Sci* 21:375-382.
51. R.A. Kammerer (1996) *Matrix Biol* 15: 555-565.
52. W.D. Kohn (1997) *J. Biol. Chem.* 272:2583-2586.
53. C. Wild (1993) *AIDS Res Hum Retroviruses*, 9:1051-1053.
54. C Wild (1994) *Proc Natl Acad Sci U S A*; 91:9770-9774.
55. United States Patent No. 4,816,567, issued March 28, 1989 to Cabilly et al.
56. United States Patent No. 5,225,539, issued July 6, 1993 to Gregory Winter.
- 10 57. United States Patent No. 5,585,089, issued December 17, 1996 to Queen et al.
58. United States Patent No. 5,693,761, issued December 2, 1997 to Queen et al.
- 15 59. PCT International Application No. PCT/US89/05857, filed December 28, 1989, published July 26, 1990, WO 90/07861.
60. Clercq, E.D. et al. (1994) *Antimicrobial Agents and Chemotherapy* 38:668-674.
61. Clercq, E.D. et al (1992) *Proc. Natl. Acad. Sci. USA* 89: 5286-5290.
- 20 62. Merrifield, R.B. (1963) *J. Am. Chem. Soc.* 85: 2149-2154.
63. Goeghegan, K.F. Stroh, J.F. (1992) *Bioconjugate Chem.* 3: 138-146.
- 25 64. Baba, M. et al. (1999) *Proc. Natl Acad Sci USA*, 96(10):5698-5703.
65. Schols D. et al (1997) *Antiviral Res*, 35(3):147-156.

Second Series of Experiments

CD4 and CCR5 mediate fusion and entry of R5 HIV-1 strains. Sulfotyrosine and other negatively charged residues in the CCR5 amino-terminal domain (Nt) are crucial for gp120 binding and viral entry. It is shown that a soluble gp120/CD4 complex specifically binds to a peptide corresponding to CCR5 Nt residues 2-18, with sulfotyrosines in positions 10 and 14. This sulfopeptide also inhibits soluble gp120/CD4 binding to cell surface CCR5 as well as infection by R5 virus. These observations prompted the further delineation of the determinants of the gp120-CCR5 Nt sulfopeptide interaction. It is shown that residues 10-18 constitute the minimal domain of the CCR5 Nt that is able to specifically interact with soluble gp120/CD4 complexes. In addition to sulfotyrosines in positions 10 and 14, negatively charged residues in positions 11 and 18 participate in this interaction. Furthermore, the CCR5 Nt binds to a CD4-induced surface on gp120 that is composed of conserved residues in the V3 loop stem and the C4 domain. Binding of gp120 to cell surface CCR5, however, is further influenced by variable residues in the crown of the V3 loop. This data suggest that gp120 docking to CCR5 is an interdependent, multi-step process involving different regions of the envelope glycoprotein and the co-receptor.

25

Entry of HIV-1 R5 isolates into target cells is mediated by the successive interaction of the envelope glycoprotein gp120 with CD4 and the CCR5 co-receptor [3]. Gp120-CD4 complex formation generates a large bonding energy that drives reordering of the gp120 core structure [22, 31, 47].

Changes in the orientation of the V1/V2 and V3 loops, as well as the bridging sheet (composed of the V1/V2 stem and C4), cooperatively create and/or expose a co-receptor binding site on gp120 [22, 37, 47]. The predicted co-receptor 5 binding surface on gp120 has a hydrophobic core surrounded by a positively charged periphery and is composed of both conserved and variable residues located in the C4 domain and V3 loop, with lesser contributions from the V1V2 stem [22, 36, 37].

10

It has been demonstrated that specific amino acids within the CCR5 amino-terminal domain (Nt, amino acids 2-31), including negatively charged and tyrosine residues, are essential for CCR5-mediated fusion and entry of R5 and R5X4 15 HIV-1 strains [5, 12, 13, 15, 35]. Farzan et al. [16] demonstrated that the CCR5 Nt undergoes both O-glycosylation and tyrosine sulfation. It is presently not known whether O-glycosylation plays a role in co-receptor function, but this possibility is suggested by observations 20 that serines in the Nt are important for viral entry. Inhibition of cellular sulfation pathways, including tyrosine sulfation, greatly decreases gp120 binding to CCR5 as well as the entry of R5 and R5X4 HIV-1 strains into target cells ([16], E.G.C. unpublished data). Post- 25 translational sulfation of the tyrosine residues in the CCR5 Nt, therefore, may critically modulate the susceptibility of target cells to HIV-1 infection *in vivo*.

It was demonstrated that a CCR5 Nt-based peptide spanning 30 residues 2-18 and containing sulfotyrosines in positions 10

and 14 specifically associates with soluble gp120/CD4 complexes containing envelope glycoproteins from R5 (JR-FL) and R5X4 (DH123) but not X4 (LAI) strains [11]. Peptides containing unmodified tyrosines or phosphotyrosines, 5 however, did not bind soluble gp120/CD4 complexes [11]. The tyrosine-sulfated CCR5 Nt therefore specifically interacts only with gp120 proteins from isolates that use this co-receptor to gain entry into target cells. Furthermore, only the CCR5 Nt-based sulfopeptide inhibits binding of soluble 10 gp120<sub>JR-FL</sub>/CD4 to intact, cell surface-expressed CCR5 and moderately blocks the entry of the R5 isolate JR-FL. The affinity of soluble gp120/CD4 for the CCR5 Nt sulfopeptide, however, is approximately 10-100-fold lower than for the native, membrane-associated co-receptor [11, 42, 46], 15 suggesting that other gp120-CCR5 contacts are required to consolidate this interaction. This concept is further supported by studies of CCR5 chimera, as well as studies with inhibitors of CCR5 co-receptor function [12, 34, 38, 32, 14].

20

A novel ELISA is reported to detect binding of sulfopeptides to soluble gp120/CD4 complexes, as well as anti-CCR5 MAbs and chemokines. ELISA and surface plasmon resonance (SPR) were used to further delineate the 25 determinants of the gp120-CCR5 Nt interaction. In order to define the minimal domain of the CCR5 Nt capable of specifically binding to soluble gp120/CD4 complexes, sulfopeptides corresponding to different regions of the Nt were analyzed. To identify the gp120 domains involved in 30 sulfopeptide binding, inhibition of gp120/CD4 complex

binding to CCR5 Nt sulfopeptides by anti-gp120 Mabs was studied. Residues in or near the epitopes of inhibitory MAbs were mutated to alanine, and the gp120 point mutants were compared for their ability to bind to CCR5 Nt sulfopeptides and cell-surface CCR5. The data suggest that a mostly conserved surface of gp120 binds to a nine-residue stretch of the CCR5 Nt, whereas more variable residues in the crown of the V3 loop may interact with a secondary binding site on CCR5.

10

#### Materials and Methods

Reagents: CD4-IgG<sub>2</sub>, soluble CD4 (sCD4), recombinant soluble gp120s from HIV-1<sub>LAI</sub> (X4), HIV-1<sub>DH123</sub> (R5X4), and HIV-1<sub>JR-FL</sub> (R5) isolates, anti-gp120 MAb PA1 (directed against the V3 loop of JR-FL) and anti-CCR5 MAbs PA8, PA10, PA11, PA12, PA14 were produced by Progenics Pharmaceuticals, Inc. (Tarrytown, NY) as described [1, 32]. MAbs 133-290, 133-192, 135-9, A32, 17b, 19b, 48d, 9284, G3-42, C11, G45-60 and 2G12 were a generous gift [26]. The small-molecule CCR5 antagonist TAK-779 was obtained as described [14].

Peptides corresponding to different segments of the CCR5 Nt were synthesized as described previously (Table 1) [11]. Sulfo- or phospho-tyrosines were incorporated in positions 10 and 14, and all peptides carried a carboxy-terminal Gly-Ala-Gly spacer preceding a biotinylated lysine. Residues were numbered according to their positions in the full length CCR5 protein.

30

Surface plasmon resonance: Binding of gp120/CD4-IgG<sub>2</sub> complex and MAbs to CCR5 Nt-based peptides was measured as previously described [11]. Briefly, streptavidin-coated sensor chips were divided into two surfaces, each with a 5 separate flow chamber. The sensor chip was conditioned and equilibrated as recommended by the manufacturer. Biotinylated peptide (400 resonance units, RU) was bound to the surface of the second chamber whereas the first chamber of the chip was used as a negative control. Gp120/CD4-IgG<sub>2</sub>, 10 complex (50 nM) was passed over the chip surface in the presence or absence of MAbs (150 mM). Surface plasmon resonance was monitored and displayed in RU as a function of time using a Biacore X. After each measurement the chip was regenerated and equilibrated as recommended by the 15 manufacturer.

Generation of gp120 alanine mutants and their binding to CD4-IgG<sub>2</sub>: Mutant gp120 proteins were generated using the QuickChange Kit from Stratagene (San-Diego, CA). Gp120<sub>JR-FL</sub>, 20 cloned into the pPPI4 expression vector [4], served as the template for site directed mutagenesis. Nucleotide sequencing was performed to ascertain the presence of the appropriate mutation in the gp120 coding sequence. 293T cells were calcium phosphate transfected with the different 25 mutant gp120 expression constructs. Supernatants containing soluble gp120 proteins were harvested and cleared of debris by centrifugation 24 hours post-transfection. Quantification of gp120 was performed by ELISA as previously described [40]). Briefly, 293T supernatants were 30 boiled for 5 minutes and denatured gp120 was captured on an

This page is not part of  
the pamphlet!

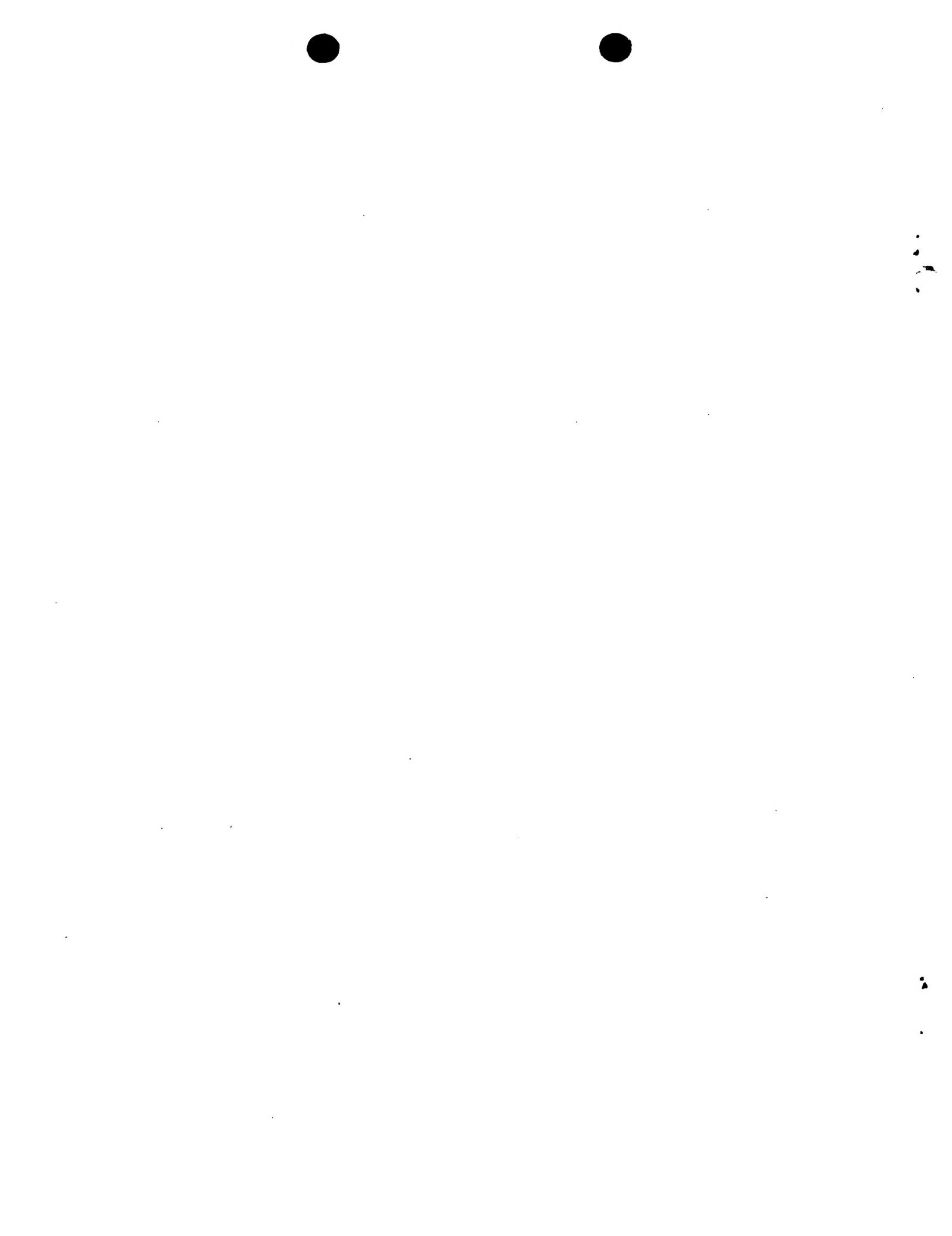
**WO 01-64710**

**2/2**

Date: 07 sep 2001

Destination: Agent

Address:



ELISA plate coated with D7324 (International Enzymes Inc. Fallbrook, CA), a MAb that recognizes a 15 residue linear epitope in the carboxy-terminal end of gp120. Captured gp120 was detected by a mixture of anti-gp120 MAbs B12 and 5 B13 [40], followed by incubation with a horseradish peroxidase-conjugated (HRP) anti-mouse IgG antibody (Amersham Pharmacia, Piscataway, N.J.). Optical density (O.D.) was measured at 450 nm using the ImmunoPure TMB Substrate kit (Pierce, Rockford, IL).

10

CD4-IgG<sub>2</sub> binding to non-denatured mutant gp120 proteins also was measured. Plates coated with D7324 were used to capture native gp120 from supernatants of transiently transfected 293T cells. CD4-IgG<sub>2</sub> (50 nM) was added to the plates and its 15 binding was detected using an HRP-conjugated goat anti-human IgG and TMB substrate as described above.

CCR5 Nt peptide ELISAs: Streptavidin-coated ELISA plates (Pierce, Rockford, IL) were blocked with D-PBS/ 5% BSA for 20 2 hours at room temperature then washed three times with assay buffer (D-PBS/ 0.5% Tween 20/ 1% Fetal Bovine Serum/ 2% BSA). Plates were then contacted with sulfo- or phosphopeptides (1  $\mu$ g/ml) for 1 hour at room temperature and washed three times with assay buffer. Mixtures of CD4-IgG<sub>2</sub> 25 (50 nM) and purified gp120 or gp120-containing supernatants in a 1:4 molar ratio were added to the plates for 1 hour at room temperature. Plates were washed three times and (HRP)-conjugated goat anti-human IgG was used to detect the presence of bound CD4-IgG<sub>2</sub>. The plates were developed using 30 the TMB substrate as described above. Gp120/CD4-IgG<sub>2</sub> binding

to the peptides was normalized for CD4-IgG<sub>2</sub> binding to the mutant gp120 proteins.

In a competition ELISA, peptides were captured onto the 5 plates as described above. Inhibitor or assay buffer was added for 1 h prior to addition of gp120<sub>JR-FL</sub>/CD4-IgG<sub>2</sub> complex (1 nM) for an additional h at room temperature. The assay was then completed as described above. Direct binding of 10 anti-CCR5 murine MAbs to the peptides was examined as described above except that MAb was substituted for gp120/CD4-IgG<sub>2</sub> complex and a goat anti-mouse HRP-coupled antibody was used for detection.

Binding of gp120/CD4-IgG<sub>2</sub> complexes to cell-surface CCR5:  
15 L1.2-CCR5<sup>+</sup> cells ( $10^6$ ) were incubated with gp120-containing supernatant (100nM) and biotinylated CD4-IgG<sub>2</sub> (50nM) for 1 hour at 37°C in assay buffer, as previously described [32]. Gp120/CD4-IgG<sub>2</sub> bound to the cells was revealed by FACS analysis of the mean fluorescence of intensity (m.f.i.) 20 after addition of streptavidin-PE (Pharmingen, San-Diego, CA.). Binding was calculated using the formula: (m.f.i. gp120 mutants)/(m.f.i. gp120 wild type) x 100% and normalized for CD4-IgG<sub>2</sub> binding to the mutant gp120 proteins.

25

Results

An ELISA to detect binding of soluble gp120/CD4 complexes to CCR5 Nt-based peptides: Surface plasmon resonance (SPR) 30 was previously used to show that gp120/sCD4 complexes

specifically interact with a peptide spanning CCR5 residues 2-18 and containing sulfotyrosines in positions 10 and 14 (2-18, Table 1). The on and off rates of complex-peptide binding were extremely rapid and could not be measured 5 precisely by SPR. The Kd was estimated to be in the  $10^{-7}$ - $10^{-8}$  range. Replacing monomeric sCD4 with tetravalent CD4-IgG<sub>2</sub>, however, lead to a dramatic shift in both on and off rates, lowering the Kd into the  $10^{-8}$ - $10^{-9}$  range (Figure 7a). This observation prompted us to develop an ELISA to directly 10 detect complex-peptide binding. Streptavidin-coated ELISA plates were used to capture biotinylated, CCR5 Nt-based peptides, and then further incubated with soluble gp120/CD4-IgG<sub>2</sub> complexes. Complex binding was detected using an HRP-conjugated goat anti-human IgG antibody.

15

Sulfopeptide 2-18 bound gp120<sub>JR-FL</sub>/CD4-IgG<sub>2</sub> with an IC<sub>50</sub> ~1nM, and gp120<sub>DH123</sub>/CD4-IgG<sub>2</sub> with an IC<sub>50</sub> ~5nM (Figure 7b). Sulfopeptide 2-18 did not measurably bind CD4-IgG<sub>2</sub> alone or in complex with either gp120<sub>LAI</sub> or V3 loop-deleted gp120<sub>JR-FL</sub> 20 (Figure 7b and data not shown). No binding was observed to an analogous CCR5 Nt phosphopeptide (2-18(P), Figure 12) by any of the gp120/CD4-IgG<sub>2</sub> complexes (Figure 7b). Identical patterns of reactivity were observed for gp120s in complex with CD4-γ2, a divalent CD4-immunoglobulin fusion protein 25 [data not shown, (1)]. However, no binding was observed for gp120 in complex with anti-gp120 MAbs 2G12 and IgG1b12, even though the latter's epitope overlaps the CD4 binding site on gp120 (data not shown). Thus the ELISA reproduces the critical biological features of cell-surface CCR5/gp120 30 interactions, including a dependence upon CCR5 tyrosine

sulfation, CD4, the V3 loop, and the coreceptor usage patterns of the parent viruses.

Using a competition ELISA, inhibition of gp120<sub>JR-FL</sub>/CD4-IgG2  
5 binding to the sulfopeptide 2-18 was enabled with the anti-CCR5 MAb PA8. However, binding of soluble gp120/CD4-IgG<sub>2</sub> complexes to the sulfopeptide was not inhibited by TAK-779, nor the CC-chemokines MIP-1 $\alpha$ , MIP-1 $\beta$  and RANTES even when used at supraphysiologic concentrations.

10

Binding of CCR5 Nt peptides to anti-CCR5 MAbs and soluble gp120/CD4: ELISA was used to test the binding of a panel of anti-CCR5 MAbs to peptides 2-18 and 2-18(P). We had previously demonstrated that MAbs PA8, PA11 and PA12 bind epitopes in the Nt, PA10 binds an epitope that spans the Nt and ECL2 and PA14 binds an epitope exclusively in ECL2 [32]. Here we show that PA8 avidly binds peptides 2-18 and 2-18(P) (Figure 2). PA10 binds avidly to 2-18 and moderately to 2-18(P). PA11 binds moderately to 2-18 and 20 weakly to 2-18(P). PA12's binding is weak for 2-18 and undetectable for 2-18(P). Finally, PA14 does not recognize either the sulfopeptide or the phosphopeptide (Figure 8). Furthermore, PA8 binds similarly to all of the CCR5 Nt-based sulfopeptides in Figure 12 (data not shown).

25

In order to more precisely delineate the minimal CCR5 Nt domain that specifically binds to soluble gp120/CD4 complexes, a panel of sulfopeptides spanning different stretches of the CCR5 Nt were synthesized (Figure 12). All 30 of the peptides carried sulfotyrosines in positions 10 and

14 since we previously showed that these are required for complex-peptide binding. Binding of gp120<sub>JR-FL</sub>/CD4-IgG<sub>2</sub> to the different sulfopeptides was tested by ELISA. Although the strongest binding was observed using longest sulfopeptide,  
5 2-18, significant binding for peptide 10-18, which demonstrated ~3-fold lower avidity was also observed (Figure 9). Peptides 8-15, 6-16 and 10-15 bound the soluble complex at least ten-fold less avidly than 2-18. (It was previously shown that a sulfopeptide consisting of residues  
10 10-14 did not bind soluble gp120/CD4 complexes.) Furthermore, the gp120<sub>JR-FL</sub>/CD4-IgG<sub>2</sub> complex only weakly bound to peptide 10-18 carrying two alanine mutations in positions 11 and 18. Previous mutagenesis studies have shown that residues Asp-11 and Glu-18 are important for  
15 fusion, entry and gp120-CCR5 binding [12, 13]. Finally, it should be noted that the same binding patterns to the different sulfopeptides were observed with soluble complexes containing gp120<sub>DH123</sub> (data not shown).

20 Inhibition of gp120/CD4 binding to CCR5 Nt sulfopeptides by anti-gp120 MAbs: In order to determine which domains of gp120 were involved in binding to CCR5 Nt-based sulfopeptides, the ability of a panel of well-characterized anti-gp120 MAbs [25] to inhibit binding of either the  
25 gp120<sub>JR-FL</sub>/CD4-IgG<sub>2</sub> or the gp120<sub>DH123</sub>/CD4-IgG<sub>2</sub> complex to the 2-18 sulfopeptide was tested. Only MAbs directed against CD4-induced (CD4i) epitopes and the V3 loop were capable of inhibiting binding of the gp120/CD4-IgG<sub>2</sub> complex to the CCR5 sulfopeptide. Inhibition of gp120<sub>JR-FL</sub> and gp120<sub>DH123</sub>  
30 binding by MAbs 17b and 48d was >90%. The anti-V3 loop MAb

-100-

19b [28], which recognizes an epitope in the crown of the V3 loop (sequence -I---G--FY-T) and is reactive with R5 strains, inhibited gp120<sub>JR-FL</sub> binding >90% and gp120<sub>DH123</sub> binding by approximately 80%. Anti-V3 loop MAb PA1, which 5 was raised against gp120<sub>JR-FL</sub> (W.C.O., unpublished results) efficiently inhibited binding of gp120<sub>JR-FL</sub> but not gp120<sub>DH123</sub>. Finally, anti-V3 loop MAb 9284 [21], which recognizes an epitope spanning residue 307 to 330 in the V3 loop of X4 strains, was unable to inhibit binding of either gp120 10 protein to the sulfopeptide. MAbs directed against other epitopes in other constant and variable regions of gp120 also had no effect on binding of the soluble complex to the peptides. Similar results were obtained when the anti-gp120 MAbs were used to inhibit soluble complex binding to cell 15 surface CCR5 (data not shown).

Binding of mutant soluble gp120/CD4 complexes to CCR5 Nt sulfopeptides: Numerous studies have shown that residues in the V3 loop determine co-receptor usage and binding [6-10, 20, 18, 20-21, 23-24, 27, 29, 33, 41-44, 46]. The crystal structure of a gp120 lacking the V1/V2 and V3 loops in complex with sCD4 and the 17b MAb further implicated a conserved, CD4i surface on gp120, adjacent to the V3 loop, in co-receptor binding [36, 37]. Single alanine mutants of 25 all of the residues near or within regions previously shown to be important for co-receptor usage were generated. These gp120 mutants were tested for their ability to bind to the CCR5-based sulfopeptide 2-18 as well as to cell surface CCR5. Binding was normalized for gp120 mutant binding to 30 CD4-IgG<sub>2</sub>. Wild-type levels of binding were observed for all

mutants except W427A, R440A and R469A, which bound CD4-IgG<sub>2</sub> with 5-10-fold lower but nonetheless measurable avidity.

Residues in both strands of the V3 loop stem were found to  
5 be involved in gp120 binding to the 2-18 sulfopeptide:  
Alanine mutants of residues R298, N301, T303, I322, D324,  
I325 and R326 were found to decrease complex binding to the  
peptide by >10-fold. Residues in the crown of the V3 loop,  
including the GPGR motif, however, had no effect on gp120  
10 binding to the sulfopeptide. C4 residues in or adjacent to  
the two C-terminal  $\beta$ -strands of the bridging sheet were  
also shown to participate in binding to the sulfopeptide:  
Alanine substitutions of R419, I420, K421, Q422 and R444  
decreased complex binding to the sulfopeptide by 5-10-fold.  
15 None of the alanine substitutions that we introduced in the  
other regions of gp120 significantly affected complex-  
peptide-interactions.

It was furthermore demonstrated that additional gp120  
20 residues are involved in complex binding to cell surface  
expressed CCR5. Alanine substitutions of residues S306,  
G310, P311, R313, F315, Y316 in the crown of the V3 loop  
decreased complex binding to CCR5 by 5-10-fold.  
Furthermore, alanine substitutions of several residues in  
25 C1, C2 and C3 also had a moderate effect on complex binding  
to CCR5. Finally, it is noted that alanine substitutions of  
R440 and R469 increased complex binding to both 2-18 and  
CCR5, whereas substitutions of E320 and W427 increased  
complex binding to CCR5 only.

Discussion

Tyrosine-sulfated CCR5 Nt peptides were studied for binding to soluble gp120/CD4 complexes as well as anti-CCR5 MAbs, CC-chemokines and TAK-779 using a novel solid-phase ELISA.

5 Inhibition of peptide-complex interactions by anti-gp120 MAbs was explored by surface plasmon resonance. These MAbs were also tested for their ability to inhibit complex binding to cell surface CCR5. In addition, a panel of gp120 point mutants were generated and then their 10 reactivity was compared with CCR5 Nt peptides and cell surface CCR5. The principal conclusions are that (1) residues 10 to 18 of the CCR5 Nt may define the minimum recognition site for gp120, (2) gp120 binding to the CCR5 Nt depends on highly conserved residues located in the C4 15 domain and the stem of the V3 loop, and (3) gp120 binding to cell surface CCR5 depends on a broader region that includes residues in the crown of the V3 loop, C1, C2 and C3. The findings suggest that distinct domains of gp120 and CCR5 bind in a multi-step fashion and raise questions 20 about the determinants of specificity of the co-receptor-gp120 interaction.

An ELISA was developed to detect complex-peptide binding based on the observation that the tetravalent gp120/CD4-IgG<sub>2</sub> 25 complex binds to CCR5 Nt sulfopeptides ten-to a hundred - fold more avidly than the monovalent gp120/sCD4 complex. Complex-sulfopeptide binding was only observed with gp120 proteins derived from R5 and R5X4, but not X4 HIV-1 strains. V3 loop deleted gp120<sub>JR-FL</sub> failed to bind to the 30 sulfopeptides. Phosphopeptides did not bind to any of the

soluble complexes. Thus, the ELISA reproduces the critical biological features of cell-surface CCR5-gp120 interactions, including a dependence upon CD4, CCR5 tyrosine sulfation, the V3 loop and the co-receptor usage 5 patterns of the parental viruses.

CCR5 Nt phosphopeptides and sulfopeptides were differentially recognized by anti-CCR5 MAbs in ELISA. PA8 possessed equal avidity for sulfated and phosphorylated 10 peptides, implying that its epitope does not include tyrosine side chains. PA10 and PA11 preferentially recognized the sulfopeptide, albeit with varying efficiencies, suggesting that sulfotyrosines participate either directly in peptide-MAb interactions or indirectly 15 by influencing epitope conformations. PA12 only interacted with the sulfopeptide and PA14 did not bind either Nt peptide. It was previously shown that both of these MAbs recognize discontinuous epitopes comprising residues in the Nt and ECL2 of CCR5. The observations now imply that ECL2 20 residues are marginal for PA12 binding and essential for PA14 binding to CCR5. Finally, binding of soluble gp120/CD4 complexes to CCR5 Nt peptides could be completed with an anti-CCR5 Mab but not with either CC-Chemokines or TAK-779, whose binding sites have been mapped to other 25 regions of CCR5 (14, 39). Both CC-chemokines and TAK-779, however, are able to compete with gp120/CD4 binding to cell surface CCR5, perhaps through steric or conformational mechanisms. (14, 42, 46). It is noted that Farzan *et al.* reported that a CCR5 Nt sulfopeptide spanning residues 1-22 30 partially blocks MIP-1 $\alpha$  binding to cell-surface CCR5 and we

attribute the discrepancy to differences in peptides and assays (17).

In order to more precisely delineate the gp120 binding site in the CCR5 Nt, an ELISA was used to test binding of soluble complexes to several CCR5 Nt sulfopeptides. 10-18 was the smallest sulfopeptide that avidly bound soluble gp120/CD4 complexes and may define the minimum docking site for gp120 on CCR5. In addition to the two sulfotyrosines in positions 10 and 14, negatively charged amino acids D11 and E18 were found to be critical for complex-peptide binding. It was concluded that a cluster of negative charges in the CCR5 Nt appears to represent the principal recognition motif for gp120, although residues 2 to 9 further contribute to binding. Similar patterns of peptide reactivity were observed for recombinant gp120s derived from HIV-1<sub>JR-FL</sub> (R5) and HIV-1<sub>DH123</sub> (R5X4), suggesting that the CCR5 Nt sulfopeptides recognize conserved structures in the envelope glycoprotein. Gp120<sub>DH123</sub>, however, bound about five-fold less than gp120<sub>JR-FL</sub> to the sulfopeptides, which probably accounts for its less efficient usage of CCR5 (13).

Anti-gp120 MAbs were tested for their ability to inhibit gp120/CD4 binding to sulfopeptides or to cell surface CCR5. A number of anti-gp120 MAbs directed against conserved and variable regions of the envelope glycoprotein were not inhibitory. Only MAbs 48d and 17b, directed against CD4i epitopes, and 19b and PA1, directed against the V3 loop, efficiently inhibited gp120 binding to the 2-18

sulfopeptide and to cell surface CCR5. The CD4i epitope was previously shown to participate in co-receptor binding and residues in the V3 loop primarily determine co-receptor specificity (36, 37). The results now suggest that these 5 regions of gp120 determine its association with the CCR5 Nt. It is noted that inhibition of peptide-complex binding by 19b, which recognizes an epitope in the V3 crown, is inconsistant with the finding by gp120 mutagenesis experiments that residues in the V3 loop crown do not 10 participate in complex-peptide binding. This leads to a conclusion that the inhibitory effect of 19b may be steric hindrance.

In order to determine more precisely the regions of gp120 15 that modulate the gp120-CCR5 interaction, the binding of a panel of gp120 point mutants to the CCR5 Nt sulfopeptide and to cell surface CCR5 was tested. The mutants were created by the introduction of single alanine substitutions near or within regions previously shown to be important for 20 the integrity of the CD4i epitope and/or CCR5 binding (36, 37). Highly conserved residues in C4 and the V3 loop stem, including for arginines and a lysine, were found to affect binding of gp120 to the CCR5 Nt sulfopeptide (Figure 13). These residues are located in two random coil segments of 25 C4 that straddle the V3 loop stem and may constitute a positively charged CCR5 Nt binding domain (22). Additional, conserved residues in the crown of the V3 loop, C1, C2 and C3 contribute to gp120 binding to cell surface CCR5 (Figure 13). It remains to be determined whether 30 these residues interact with other extracellular domains of

-106-

CCR5 or whether they influence the conformation of C4 and the V3 loop stem in a way that is only relevant in the context of gp120/CD4 binding to cell surface CCR5. It is unlikely that these residues also interact with the Nt in 5 the context of cell surface CCR5 because they are relatively distal from the C4 and V3 residues that were implicated in sulfopeptide binding (22).

To date, several lines of evidence suggest that gp120 binds 10 to more than one region of the CCR5 co-receptor: (1) the affinity of gp120s/CD4 for the CCR5 Nt sulfopeptide is approximately 10-100-fold lower than for the native, membrane-associated co-receptor (11, 42, 46), (2) co-receptor chimera studies implicate the extracellular loops 15 in viral fusion and entry (2, 12, 34, 38) and (3) inhibitors of CCR5 co-receptor function such as Mabs 2D7 and PA14, as well as TAK-779 do not bind to the CCR5 Nt yet block gp120/CD4 binding to CCR5 (14, 32). The present findings could be interpreted to support a distributed 20 model for gp120-CCR5 interactions that mirrors the two-site paradigm proposed for the interaction of certain chemokines with their receptors (30, 45). In this model, binding is initially driven by electrostatic interactions between negatively charged residues in the receptor Nt and basic 25 surfaces on the chemokine ligand. This binding serves to orient the ligand and promote its interactions with other portions of the chemokine receptor. The V3 loop crown may form initial electrostatic interactions with the extracellular loops of CCR5, which would allow the CCR5 Nt 30 to bind to a conserved region of gp120 comprising residues

in C4 and the V3 loop stem. Alternatively, the CCR5 Nt could first bind the C4/V3 stem domain, which would then promote an interaction of the V3 loop with some other region of CCR5. All of these interactions involve 5 additional gp120 residues that we have yet to identify. The role of the putative second interaction is unclear but it may further stabilize the gp120-CCR5 interaction, optimally orienting the fusion apparatus, or triggering gp41 conformational changes that are required for fusion.

10

The findings present us with a seeming paradox wherein nine residues of the CCR5 Nt confer specificity on the CCR5-gp120 interaction by binding to gp120 residues that are highly conserved among clade B isolates, regardless of 15 their co-receptor usage. However, although the C4 and V3 stem residues themselves are conserved, their precise placement may differ for R5 and X4 viruses. Clearly, relatively minor differences in the orientation, exposure or relative positioning of these widely separated residues 20 could abrogate binding to a short peptide but not a MAb (e.g., 17b) possessing a larger, more distributed binding site (37). In addition, more variable amino acids (e.g., 324) within or outside the C4/V3 loop stem may contribute to the specificity of the gp120-Nt interaction, and we 25 showed that residues N279, R313, P369 and R444 participate in gp120/CD4 binding to cell surface CCR5 but not to CCR5 Nt sulfopeptides. Future studies employing additional gp120 mutants together with CCR5 mutants and CXCR4-based sulfopeptides will shed light on the specificity 30 determinants of the gp120-co-receptor interaction.

-108-

CD4 and CCR5 mediate fusion and entry of R5 HIV-1 strains. Sulfotyrosines and negatively charged residues in the CCR5 Nt are crucial for binding of gp120 and viral entry.

5 Soluble gp120/CD4 complexes specifically bind to CCR5 Nt peptides containing sulfotyrosines in positions 10 and 14. CCR5 Nt sulfotyrosines inhibit gp120/CD4 binding to CCR5 as well as viral entry. Residues in the V3 loop and the C4 region of gp120 compose a binding site for the CCR5 amino 10 terminal domain. Residues 10-18 of the CCR5 Nt constitute a minimal binding domain for gp120: sulfotyrosines Y10 and Y14 and negatively charged residues D11 and E18 are important for binding. The CCR5 Nt terminal binding site on gp120 is composed mostly of residues in the V3 loop stem 15 and the C4 domain.

References for the Second Series of Experiments

20 1. Allaway, G P., 1995. Expression and characterization of CD4-IgG2, a novel heterotetramer that neutralizes primary HIV type 1 isolates. Aids Res Hum Retroviruses. 11:533-539.

25 2. Atchison, R.E., 1996. Multiple extracellular elements CCR5 and HIV-1 entry: dissociation from response to chemokines. Science. 274:1924-1926.

30 3. Berger, E.A., 1999. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and

disease. *Annu Rev Immunol.* 17:657-700.

4. Binley, J.M., 2000. A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated structure. *J Virol.* 74:627-643.
- 10 5. Blanpain, C., 1999. Multiple charged and aromatic residues in CCR5 amino-terminal domain are involved in high affinity binding of both chemokines and HIV-1 env protein. *J. Biol. Chem.* 274:34719-34727.
- 15 6. Chan, S. Y., 1999. V3 recombinants indicate a central role for CCR5 as a coreceptor in tissue infection by human immunodeficiency virus type 1. *J Virol.* 73:2350-2358.
- 20 7. Chavda, S.C., 1994. Molecular determinants of the V3 loop of human immunodeficiency virus type 1 glycoprotein gp120 responsible for controlling cell tropism. *J Gen Virol.* 75:3249-3253.
- 25 8. Chesebro, B., 1992. Macrophage-tropic human immunodeficiency virus isolates from different patients exhibit unusual V3 envelope sequence homogeneity in comparison with T-cell-tropic isolates: definition of critical amino acids involved in cell tropism. *J Virol.* 66:6547-6554.

-110-

9. Cho, M. W., 1998. Identification of determinants on a dualtropic human immunodeficiency virus type 1 envelope glycoprotein that confer usage of CXCR4. *J Virol.* 72:2509-2515.

5

10. Cocchi, F., 1996. The V3 domain of the HIV-1 gp120 envelope glycoprotein is critical for chemokine-mediated blockade of infection. *Nat Med.* 2:1244-1247.

10 11. Cormier, E. G., 2000. Specific interaction of CCR5 amino-terminal domain peptides containing sulfo-tyrosines with HIV-1 envelope glycoprotein gp120. *Proc. Natl. Acad. Sci. USA.* 97:5762-5767.

15 12. Doranz, B. J., 1997. Two distinct CCR5 domains can mediate coreceptor usage by human immunodeficiency virus type 1. *J Virol.* 71:6305-6314.

20 13. Dragic, T., 1998. Amino-terminal substitutions in the CCR5 coreceptor impair gp120 binding and human immunodeficiency virus type 1 entry. *J Virol.* 72:279-285.

25 14. Dragic, T., 2000. A binding pocket for a small molecule inhibitor of HIV-1 entry within the transmembrane helices of CCR5. *Proc Natl Acad Sci U S A.* 97:5639-5644.

15. Farzan, M., 1998. A tyrosine-rich region in the N terminus of CCR5 is important for human

30

immunodeficiency virus type 1 entry and mediates an association between gp120 and CCR5. *J Virol.* 72:1160-1164.

5 16. Farzan, M., 1999. Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry. *Cell.* 96:667-676.

10 17. Farzan, M., 2000. A tyrosine-sulfated peptide based on the N-terminus of CCR5 interacts with a CD4-enhanced epitope of the HIV-1 gp120 envelope glycoprotein and inhibits HIV-1 entry. *J Biol Chem:*in press.

15 18. Harrove, G., 1995. Amino acid substitutions in the V3 loop are responsible for adaptation to growth in transformed T-cell lines of a primary human immunodeficiency virus type 1. *Virology.* 210:490-494.

20 19. Hung, C. S., 1999. Analysis of the critical domain in the V3 loop of human immunodeficiency virus type 1 gp 120 involved in CCR5 utilization. *J Virol.* 73:8216-8226.

25 20. Hwang, S. S., 1991. Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. *Science.* 253:71-74.

30 21. Ivanoff, L., 1991. Alteration of HIV-1 infectivity and neutralization by a single amino acid replacement in the V3 loop domain. *Aids Res. Hum. Retroviruses.*

-112-

7:595-603.

22. Kwong, P. D., 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*. 393:648-659.

5

23. Kwong, P. D., 2000. Oligomeric modeling and electrostatic analysis of the gp120 envelope glycoprotein of human immunodeficiency virus. *J Virol*. 74:1961-1972.

10

24. Menzo, S., 1998. Rare mutations in a domain crucial for V3-loop structure prevail in replicating HIV from long-term non-progressors. *Aids*. 12:985-997.

15

25. Moore, J., 1994. Probing the structure of the surface glycoprotein gp120 of human immunodeficiency virus type 1 with a panel of monoclonal antibodies. *J. Virol*. 68:469-484.

20

26. Moore, J., 1996. Antibody cross-competition analysis of the human immunodeficiency virus type 1 exterior envelope glycoprotein. *J. Virol*. 70:1863-1872.

25

27. Moore, J.P., 1991. The role of the V3 domain of gp120 in HIV infection. *AIDS*. 5(suppl.2):S21-S33.

30

28. Moore, J.P., 1995. A human monoclonal antibody to a complex epitope in the V3 region of human immunodeficiency virus type 1 has broad reactivity within and outside clade B. *J. Virol*. 69:122-130.

29. Morris, J. F., 1994. Effect of a single amino acid substitution in the V3 domain of the human immunodeficiency virus type 1: generation of revertant viruses to overcome defects in infectivity in specific cell types. *J Virol.* 68:8380-8385.

5

30. Moser, B., 1993. Interleukin-8 antagonists generated by N-terminal modification. *J Biol Chem.* 268:7125-7128.

10

31. Myszka, D. G., 2000. Energetics of the HIV gp120-CD4 binding reaction. *Proc Natl Acad Sci U S A.* 97:9026-9031.

15 32. Olsen, W. C., 1999. Differential inhibition of human immunodeficiency virus type 1 fusion, gp120 binding, and CC-chemokine activity by monoclonal antibodies to CCR5. *J. Virol.* 73:4145-4155.

20 33. Page, K., 1992. Analysis of mutations in the V3 domain of gp160 that affect fusion and infectivity. *J Virol.* 66:524-533.

25 34. Picard, L., 1997. Multiple extracellular domains of CCR-5 contribute to human immunodeficiency virus type 1 entry and fusion. *J Virol.* 71:5003-5011.

30 35. Rabut, G. E., 1998. Alanine substitutions of polar and nonpolar residues in the amino-terminal domain of CCR5 differently impair entry of macrophage- and dualtropic

isolates of human immunodeficiency virus type 1. *J. Virol.* 72:3464-3468.

36. Rizzuto, C., 2000. Fine definition of a conserved  
5 CCR5-binding region on the human immunodeficiency  
virus type 1 glycoprotein 120. *AIDS Res Hum  
Retroviruses.* 16:741-749.

37. Rizzuto, C. D., 1998. A conserved HIV gp120  
10 glycoprotein structure involved in chemokine receptor  
binding *Science.* 280:1949-1953.

38. Rucker, J., 1996. Regions in beta-chemokine receptors  
15 CCR5 and CCR2b that determine HIV-1 cofactor  
specificity. *Cell.* 87:437-446.

39. Samson, M., 1997. The second extracellular loop of  
CCR5 is the major determinant of ligand specificity. *J  
Biol Chem.* 272:24934-24941.

20 40. Sanders, R. W., 2000. Variable-loop-deleted variants  
of the human immunodeficiency virus type 1 envelope  
glycoprotein can be stabilized by an intermolecular  
disulfide bond between the gp120 and gp41 subunits.  
25 *J. Virol.* 74:5091-5100.

41. Shioda, T., 1992. Small amino acid changes in the V3  
hypervariable region of gp120 can affect the T-cell-  
line and macrophage tropism of human immunodeficiency  
30 virus type 1. *Proc Natl Acad Sci U S A.* 89:9434-9438.

42. Trkola, A., 1996. CD4-dependent, antibody sensitive interactions between HIV-1 and its coreceptor CCR5. *Nature*. 384:184-186.

5 43. Verrier, F., 1999. Role of HIV type 1 glycoprotein 120 V3 loop in determining coreceptor usage. *AIDS Res Hum Retroviruses*. 15:731-743.

10 44. Wang, W. K., 1999. Hypervariable region 3 residues of HIV type 1 gp120 involved in CCR5 coreceptor utilization: therapeutic and prophylactic implications. *Proc Natl Acad Sci U S A*. 96:4558-4562.

15 45. Wells, T. N., 1996. Selectivity and antagonism of chemokine receptors. *J Leukoc Biol*. 59:53-60.

20 46. Wu, L., 1996. CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. *Nature*. 384:179-183.

47. Wyatt, R., 1998. The HIV-1 envelope glycoproteins: fusogens, antigens and immunogens. *Science*. 280:1884-1888.

What is claimed:

1. A compound comprising the structure:



wherein each T represents a threonine, each S  
5 represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine;

10 wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal  
15 direction;

wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal  
20 direction;

25 wherein  $\theta$  represents an amino group or an acetylated amino group; wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group;

wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined together by peptide bonds;

further provided that at least two tyrosines in the compound are sulfated.

2. The compound of claim 1, wherein  $\beta$  represents less than 17 amino acids.
3. The compound of claim 1, wherein  $\beta$  represents less than 16 amino acids.
4. The compound of claim 1, wherein  $\beta$  represents less than 15 amino acids.
- 10 5. The compound of claim 1, wherein  $\beta$  represents less than 14 amino acids.
6. The compound of claim 1, wherein  $\beta$  represents less than 13 amino acids.
- 15 7. The compound of claim 1, wherein  $\beta$  represents less than 12 amino acids.
8. The compound of claim 1, wherein  $\beta$  represents less than 11 amino acids.
- 20 9. The compound of claim 1, wherein  $\beta$  represents less than 10 amino acids.
- 25 10. The compound of claim 1, wherein  $\beta$  represents less than 9 amino acids.
11. The compound of claim 1, wherein  $\beta$  represents less than 8 amino acids.

12. The compound of claim 1, wherein  $\beta$  represents less than 7 amino acids.
13. The compound of claim 1, wherein  $\beta$  represents less than 6 amino acids.
14. The compound of claim 1, wherein  $\beta$  represents less than 5 amino acids.
- 10 15. The compound of claim 1, wherein  $\beta$  represents less than 4 amino acids.
16. The compound of claim 1, wherein  $\beta$  represents less than 3 amino acids.
- 15 17. The compound of claim 1, wherein  $\beta$  represents less than 2 amino acids.
18. The compound of claim 1, wherein  $\beta$  represents less than 1 amino acid.
- 20 19. The compound of claim 1, wherein  $\alpha$  represents less than 9 amino acids.
- 25 20. The compound of claim 1, wherein  $\alpha$  represents less than 8 amino acids.
21. The compound of claim 1, wherein  $\alpha$  represents less than 7 amino acids.

22. The compound of claim 1, wherein  $\alpha$  represents less than 6 amino acids.
23. The compound of claim 1, wherein  $\alpha$  represents less than 5 amino acids.
24. The compound of claim 1, wherein  $\alpha$  represents less than 4 amino acids.
- 10 25. The compound of claim 1, wherein  $\alpha$  represents less than 3 amino acids.
26. The compound of claim 1, wherein  $\alpha$  represents less than 2 amino acids.
- 15 27. The compound of claim 1, wherein  $\alpha$  represents less than 1 amino acid.
28. A composition comprising the compound of claim 1 and a detectable marker attached thereto.
- 20 29. The composition of claim 28, wherein the detectable marker is biotin.
- 25 30. The composition of claim 28, wherein the detectable marker is attached at the C-terminus of the compound.
31. A composition which comprises a carrier and an amount of the compound of claim 1 effective to inhibit binding of HIV-1 to a CCR5 receptor on the surface of

a CD4+ cell.

32. A method of inhibiting human immunodeficiency virus infection of a CD4+ cell which also carries a CCR5 receptor on its surface which comprises contacting the CD4+ cell with an amount of the compound of claim 1 effective to inhibit binding of human immunodeficiency virus to the CCR5 receptor so as to thereby inhibit human immunodeficiency virus infection of the CD4+ cell.

33. The method of claim 32, wherein the CD4+ cell is present in a subject and the contacting is effected by administering the compound to the subject.

34. A method of preventing CD4+ cells of a subject from becoming infected with human immunodeficiency virus which comprises administering to the subject an amount of the compound of claim 1 effective to inhibit binding of human immunodeficiency virus to CCR5 receptors on the surface of the CD4+ cells so as to thereby prevent the subject's CD4+ cells from becoming infected with human immunodeficiency virus.

35. A method of treating a subject whose CD4+ cells are infected with human immunodeficiency virus which comprises administering to the subject an amount of the compound of claim 1 effective to inhibit binding of human immunodeficiency virus to CCR5 receptors on the surface of the subject's CD4+ cells so as to

thereby treat the subject.

36. The method of any one of claims 33-35, wherein the compound is administered by aerosol, intravenous, oral or topical route.  
5
37. The method of claim 33 or 35, wherein the subject is infected with HIV-1 prior to administering the compound to the subject.  
10
38. The method of claim 33 or 34, wherein the subject is not infected with HIV-1 prior to administering the compound to the subject.  
15
39. The method of claim 38, wherein the subject is not infected with, but has been exposed to, human immunodeficiency virus.  
20
40. The method of any one of claims 33-35, wherein the effective amount of the compound comprises from about 1.0 ng/kg to about 100 mg/kg body weight of the subject.  
25
41. The method of claim 40, wherein the effective amount of the compound comprises from about 100 ng/kg to about 50 mg/kg body weight of the subject.  
30
42. The method of claim 41, wherein the effective amount of the compound comprises from about 1  $\mu$ g/kg to about 10 mg/kg body weight of the subject.

43. The method of claim 42, wherein the effective amount of the compound comprises from about 100  $\mu\text{g}/\text{kg}$  to about 1 mg/kg body weight of the subject.

5 44. The method of any one of claims 33-35, wherein the subject is a human being.

45. A method of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor which 10 comprises:

- (a) immobilizing the compound of claim 1 on a solid support;
- (b) contacting the immobilized compound from step (a) with sufficient detectable CCR5 ligand to 15 saturate all binding sites for the CCR5 ligand on the immobilized compound under conditions permitting binding of the CCR5 ligand to the immobilized compound so as to form a complex;
- (c) removing any unbound CCR5 ligand;
- (d) contacting the complex from step (b) with the agent; and
- (e) detecting whether any CCR5 ligand is displaced from the complex, wherein displacement of detectable CCR5 ligand from the complex indicates 25 that the agent binds to the compound so as to thereby identify the agent as one which inhibits binding of the CCR5 ligand to the CCR5 receptor.

46. A method of identifying an agent which inhibits 30 binding of a CCR5 ligand to a CCR5 receptor which

comprises:

- (a) contacting the compound of claim 1 with sufficient detectable CCR5 ligand to saturate all binding sites for the CCR5 ligand on the compound under conditions permitting binding of the CCR5 ligand to the compound so as to form a complex;
- 5 (b) removing any unbound CCR5 ligand;
- (c) measuring the amount of CCR5 ligand which is bound to the compound in the complex;
- 10 (d) contacting the complex from step (a) with the agent so as to displace CCR5 ligand from the complex;
- (e) measuring the amount of CCR5 ligand which is bound to the compound in the presence of the agent; and
- 15 (f) comparing the amount of CCR5 ligand bound to the compound in step (e) with the amount measured in step (c), wherein a reduced amount measured in step (e) indicates that the agent binds to the compound so as to thereby identify the agent as one which inhibits binding of the CCR5 ligand to the CCR5 receptor.

20 47. A method of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor which comprises:

- (a) immobilizing the compound of claim 1 on a solid support;
- 25 (b) contacting the immobilized compound from step (a) with the agent and detectable CCR5 ligand under

conditions permitting binding of the CCR5 ligand to the immobilized compound so as to form a complex;

- (c) removing any unbound CCR5 ligand;
- 5 (d) measuring the amount of detectable CCR5 ligand which is bound to the immobilized compound in the complex;
- (e) measuring the amount of detectable CCR5 ligand which binds to the immobilized compound in the absence of the agent;
- 10 (f) comparing the amount of CCR5 ligand which is bound to the immobilized compound in step (e) with the amount measured in step (d), wherein a reduced amount measured in step (d) indicates that the agent binds to the compound or CCR5 ligand so as to thereby identify the agent as one which inhibits binding of the CCR5 ligand to the CCR5 receptor.

20 48. The method of claim 47, wherein the amount of the detectable ligand in step (a) and step (e) is sufficient to saturate all binding sites for the CCR5 ligand on the compound.

25 49. A method of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor which comprises:

- (a) contacting the compound of claim 1 with the agent and detectable CCR5 ligand under conditions permitting binding of the CCR5 ligand to the

compound so as to form a complex;

(b) removing any unbound CCR5 ligand;

(c) measuring the amount of detectable CCR5 ligand which is bound to the compound in the complex;

5 (d) measuring the amount of detectable CCR5 ligand which binds to the compound in the absence of the agent;

(e) comparing the amount of CCR5 ligand which is bound to the compound in step (c) with the amount measured in step (d), wherein a reduced amount measured in step (c) indicates that the agent binds to the compound or CCR5 ligand so as to thereby identify the agent as one which inhibits binding of the CCR5 ligand to the CCR5 receptor.

15

50. The method of claim 49, wherein the amount of the detectable ligand in step (a) and step (d) is sufficient to saturate all binding sites for the CCR5 ligand on the compound.

20

51. The method of any one of claims 45-50, wherein the detectable CCR5 ligand is labeled with a detectable marker.

25

52. A method of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor which comprises:

a) immobilizing the compound of claim 1 on a solid support;

30 b) contacting the immobilized compound from step a)

with the agent dissolved or suspended in a known vehicle and measuring the binding signal generated by such contact;

- 5 c) contacting the immobilized compound from step a) with the known vehicle in the absence of the compound and measuring the binding signal generated by such contact;
- d) comparing the binding signal measured in step b) with the binding signal measured in step c), wherein an increased amount measured in step b) indicates that the agent binds to the compound so as to thereby identify the agent as one which binds to the CCR5 receptor.

15 53. The method of claim 52, wherein the solid support is a surface plasmon resonance sensor chip.

54. The method of claim 52 or 53, wherein the binding signal is measured by surface plasmon resonance.

20

- 55. A method of obtaining a composition which comprises:
  - (a) identifying a compound which inhibits binding of a CCR5 ligand to a CCR5 receptor according to the method of any one of claims 45-50 and 52; and
  - (b) admixing the compound so identified or a homolog or derivative thereof with a carrier.

25 56. The method of any one of claims 45-50 and 52, wherein the CCR5 ligand is a complex comprising an HIV-1 envelope glycoprotein and a CD4-based protein.

57. The method of claim 56, wherein the HIV-1 envelope glycoprotein is gp120, gp140 or gp160.

5 58. The method of claim 56, wherein the CD4-based protein is soluble CD4 or CD4-IgG2.

59. The method of any one of claims 45-50 and 52, wherein the CCR5 ligand is a chemokine.

10 60. The method of claim 59, wherein the chemokine is RANTES, MIP-1 $\alpha$  or MIP-1 $\beta$ .

61. The method of any one of claims 45-50 and 52, wherein the CCR5 ligand is an antibody.

15 62. The method of claim 61, wherein the antibody is selected from the group consisting of PA8 (ATCC Accession No. HB-12605), PA10 (ATCC Accession No. 12607), PA11 (ATCC Accession No. HB-12608), PA12 (ATCC Accession No. HB-12609).

20 63. The method of claim 45 or 47, wherein the solid support is a microtiter plate well, a bead or surface plasmon resonance sensor chip.

25 64. A compound having the structure:  
$$\Delta - (\alpha YDINYYTSE\beta\lambda)_n$$
wherein each T represents a threonine, each S represents a serine, each E represents a glutamic

30

acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine;

5 wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino 10 terminal direction;

15 wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction;

wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group;

20 wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined together by peptide bonds,

further provided that at least two tyrosines in the compound are sulfated,

25 wherein  $n$  is an integer from 1 to 8,  $\Delta$  is a polymer, and the solid line represents up to 8 linkers which attach the structure in parentheses to  $\Delta$ .

65. A compound having the structure:

$(\theta\alpha YDINYYTSE\beta)_n - \Delta$

30 wherein each T represents a threonine, each S

represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine;

5 wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the 10 I at position 9 and extending therefrom in the amino terminal direction;

15 wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined together by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction;

20 wherein  $\theta$  represents an amino group or an acetylated amino group;

wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined together by peptide bonds,

further provided that at least two tyrosines in the compound are sulfated,

25 wherein  $\pi$  is an integer from 1 to 8,  $\Delta$  is a polymer, and the solid line represents up to 8 linkers which attach the structure in parentheses to  $\Delta$ .

66. The compound of claim 64 or 65, wherein the polymer is  
30 selected from the group consisting of a linear lysine

polymer, a branched lysine polymer, a linear arginine polymer, a branched arginine polymer, polyethylene glycol, a linear acetylated lysine polymer, a branched acetylated lysine polymer, a linear chloroacetylated lysine polymer and a branched chloroacetylated lysine polymer.

5 67. The compound of claim 1, wherein the compound is a peptide which comprises consecutive amino acids having  
10 the sequence YDINYYTSE.

15 68. The compound of claim 67, wherein the tyrosines at positions 1 and 5 of the sequence YDINYYTSE are sulfated.

69. A compound comprising the structure:

$\Theta\alpha Y D \Pi \Pi Y \Pi \Pi E \beta \lambda$

wherein each E represents a glutamic acid, and each Y represents a tyrosine;

20 wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position  
25 9 and extending therefrom in the amino terminal direction;

wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order  
30 and have a sequence identical to the sequence set

forth in SEQ ID NO: 1 beginning with the P at position 19 and extending therefrom in the carboxy terminal direction;

5 wherein  $\theta$  represents an amino group or an acetylated amino group; wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group;

wherein  $\Pi$  represents any amino acid,

wherein all of  $\alpha$ , Y, D,  $\Pi$ ,  $\Pi$ , Y,  $\Pi$ ,  $\Pi$ ,  $\Pi$ ,  $\Pi$ , E and  $\beta$  are joined together by peptide bonds;

10 further provided that at least two tyrosines in the compound are sulfated.

70. The compound of claim 69, wherein the compound is a peptide which comprises consecutive amino acids have the sequence YDППУПППЕ.

71. The compound of claim 70, wherein the tyrosines at positions 1 and 5 of the sequence YDΠΠΥΠΠΠΕ are sulfated.

20

72. A compound comprising the structure:

ΘΑΥΔΙΝΥΤΣΕΒΛ

wherein each T represents a threonine, each S represents a serine, each E represents a glutamic acid, each Y represents a tyrosine; each D represents an aspartic acid, each I represents an isoleucine; and each N represents an asparagine;

wherein  $\alpha$  represents from 0 to 9 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order

and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the I at position 9 and extending therefrom in the amino terminal direction;

5 wherein  $\beta$  represents from 0 to 13 amino acids, with the proviso that if there are more than 2 amino acids, they are joined by peptide bonds in consecutive order and have a sequence identical to the sequence set forth in SEQ ID NO: 1 beginning with the P at position 10 19 and extending therefrom in the carboxy terminal direction;

wherein  $\theta$  represents an amino group or an acetylated amino group; wherein  $\lambda$  represents a carboxyl group or an amidated carboxyl group;

15 wherein all of  $\alpha, Y, D, I, N, Y, Y, T, S, E$  and  $\beta$  are joined together by peptide bonds;

further provided that at least two tyrosines in the compound are sulfated,

wherein any amino acid except for the Y at position 1, 20 D at position 2, Y at position 5 and E at position 9 may be replaced with a homologous amino acid.

73. The compound of claim 72, wherein any I amino acid residue is be replaced with a G,A,V or L amino acid residue.

74. The compound of claim 72, wherein any N amino acid residue is replaced with a Q amino acid residue.

30 75. The compound of claim 72, wherein any Y amino acid

residue is replaced with an F or W amino acid residue.

76. The compound of claim 72, wherein any T amino acid residue is replaced with an S amino acid residue.

5

77. The compound of claim 72, wherein any S is replaced with a T amino acid residue.

78. The compound of claim 72, wherein any C is replaced  
10 with an M, S, T, A, G, N, or Q amino acid residue.

15

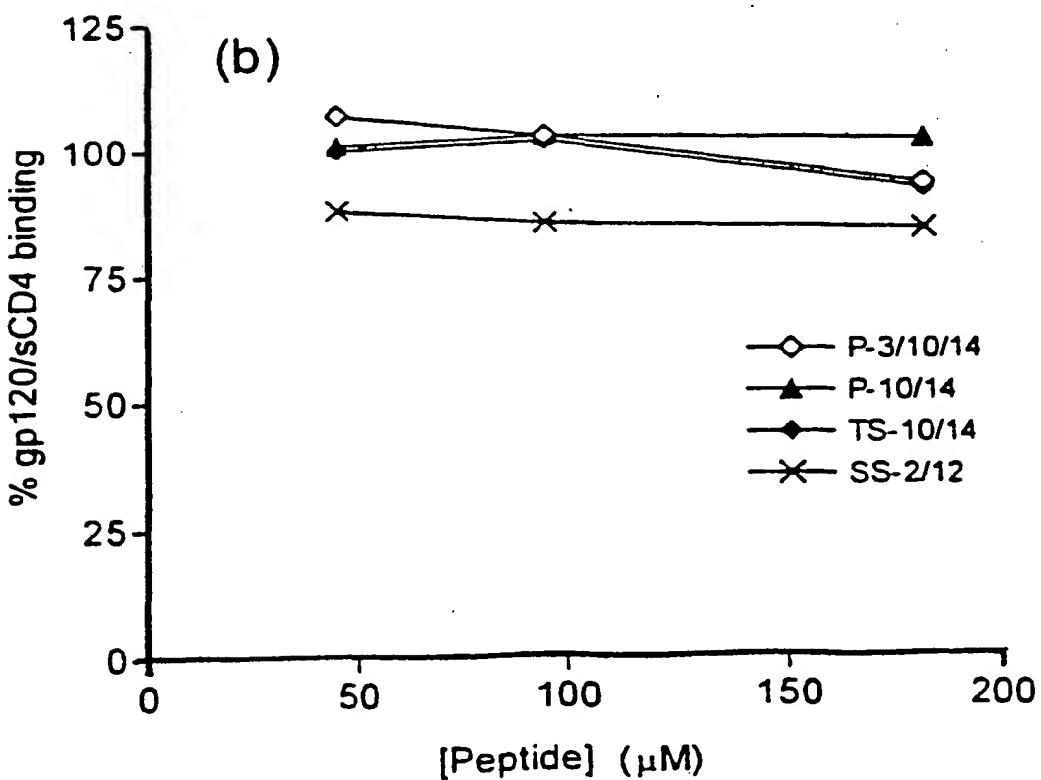
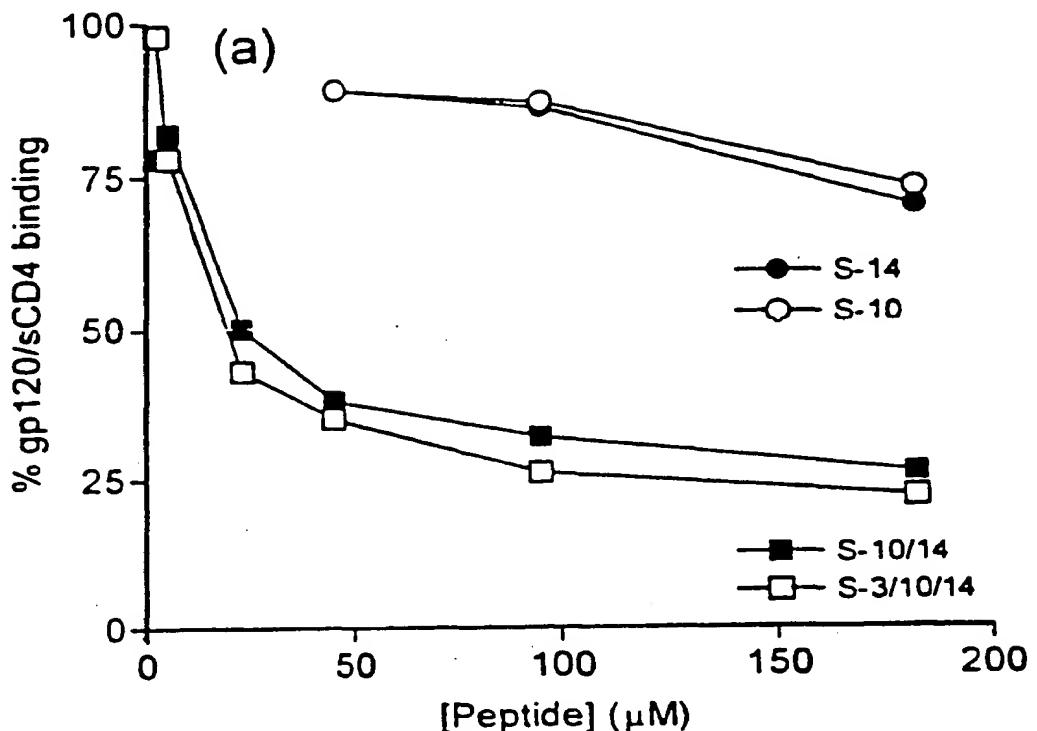
20

25

30

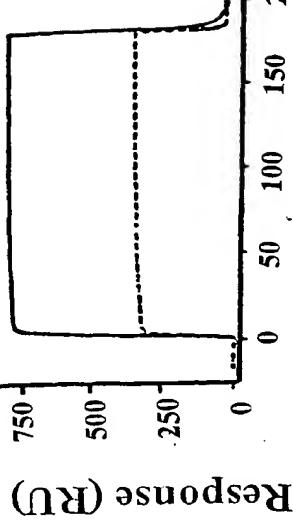


FIGURE 1

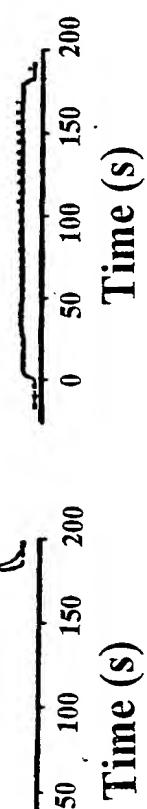


2/13

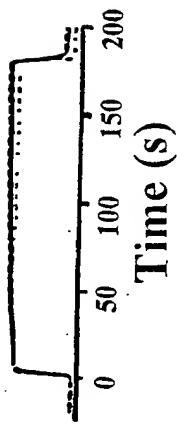
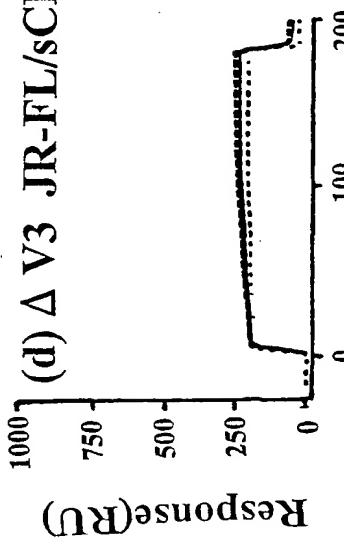
FIGURE 2 (a) JR-FL/sCD4



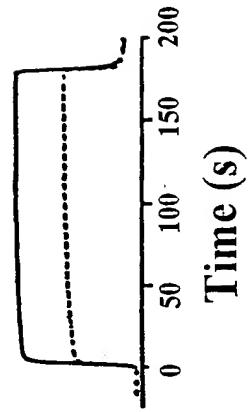
(b) JR-FL



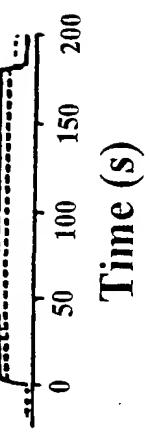
(c) sCD4

(d)  $\Delta$  V3 JR-FL/sCD4

(e) DH123/sCD4



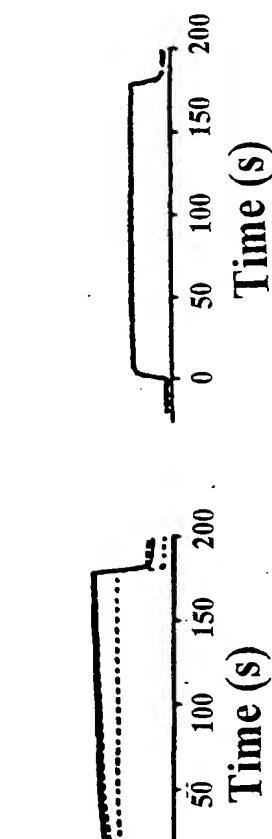
(f) DH123



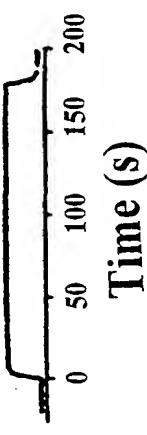
(g) LAI



(h) LAI

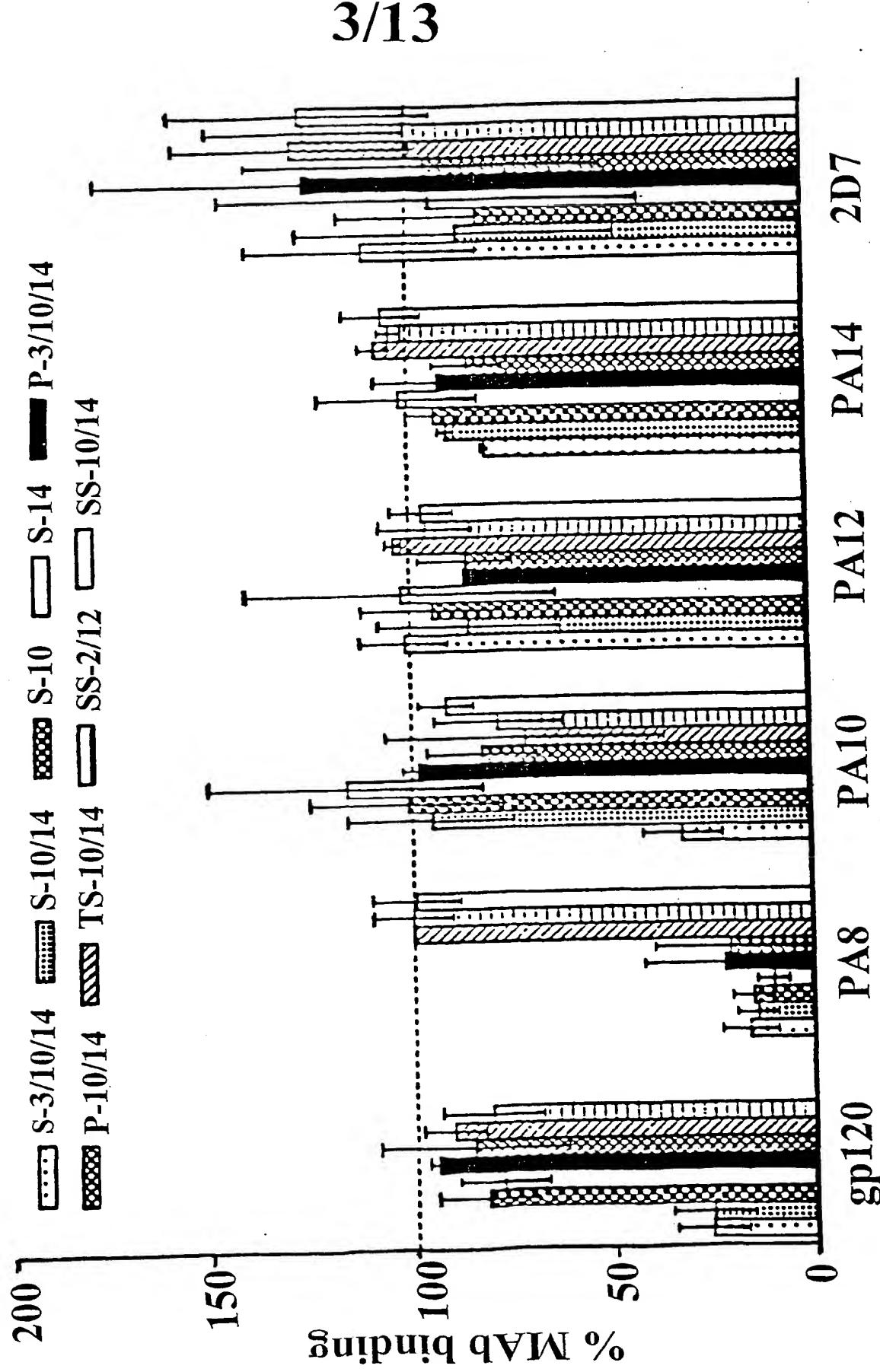


— bS-10/14  
- - - bP-10/14  
--- Empty Cell



— bS-10/14  
- - - bP-10/14  
--- Empty Cell

FIGURE 3



4/13

FIGURE 4

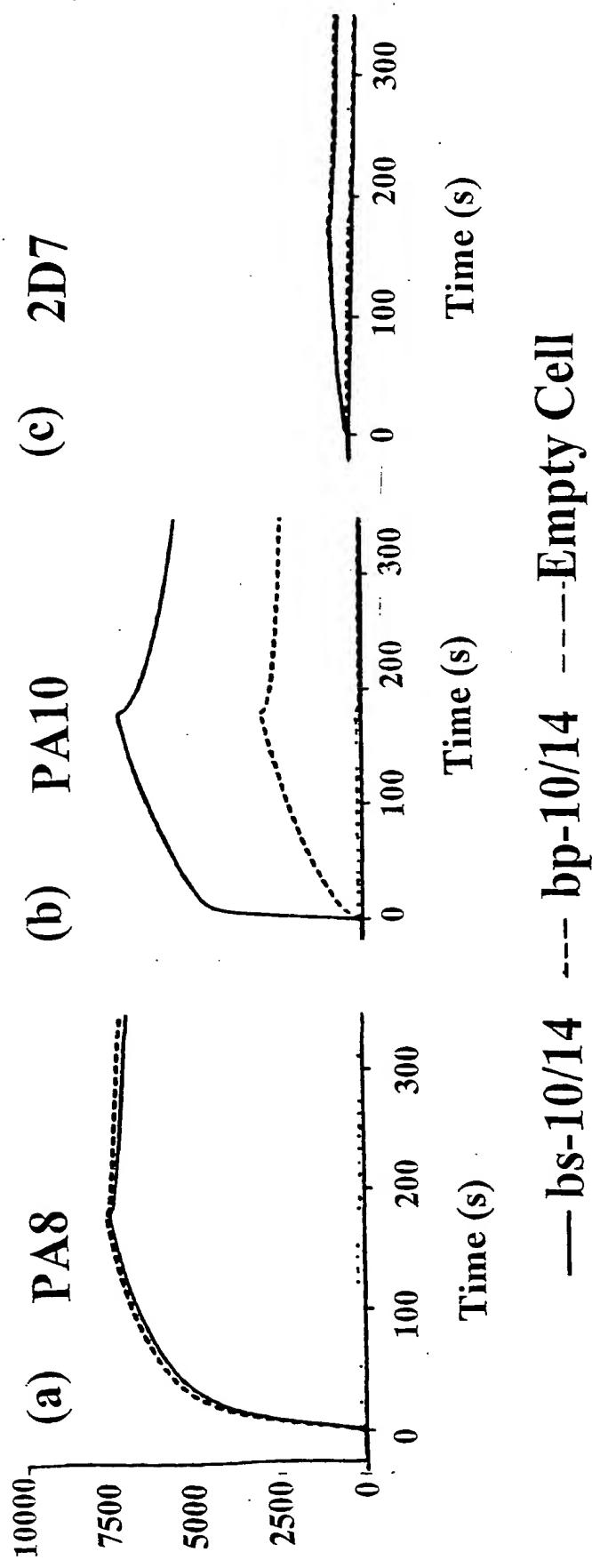


FIGURE 5

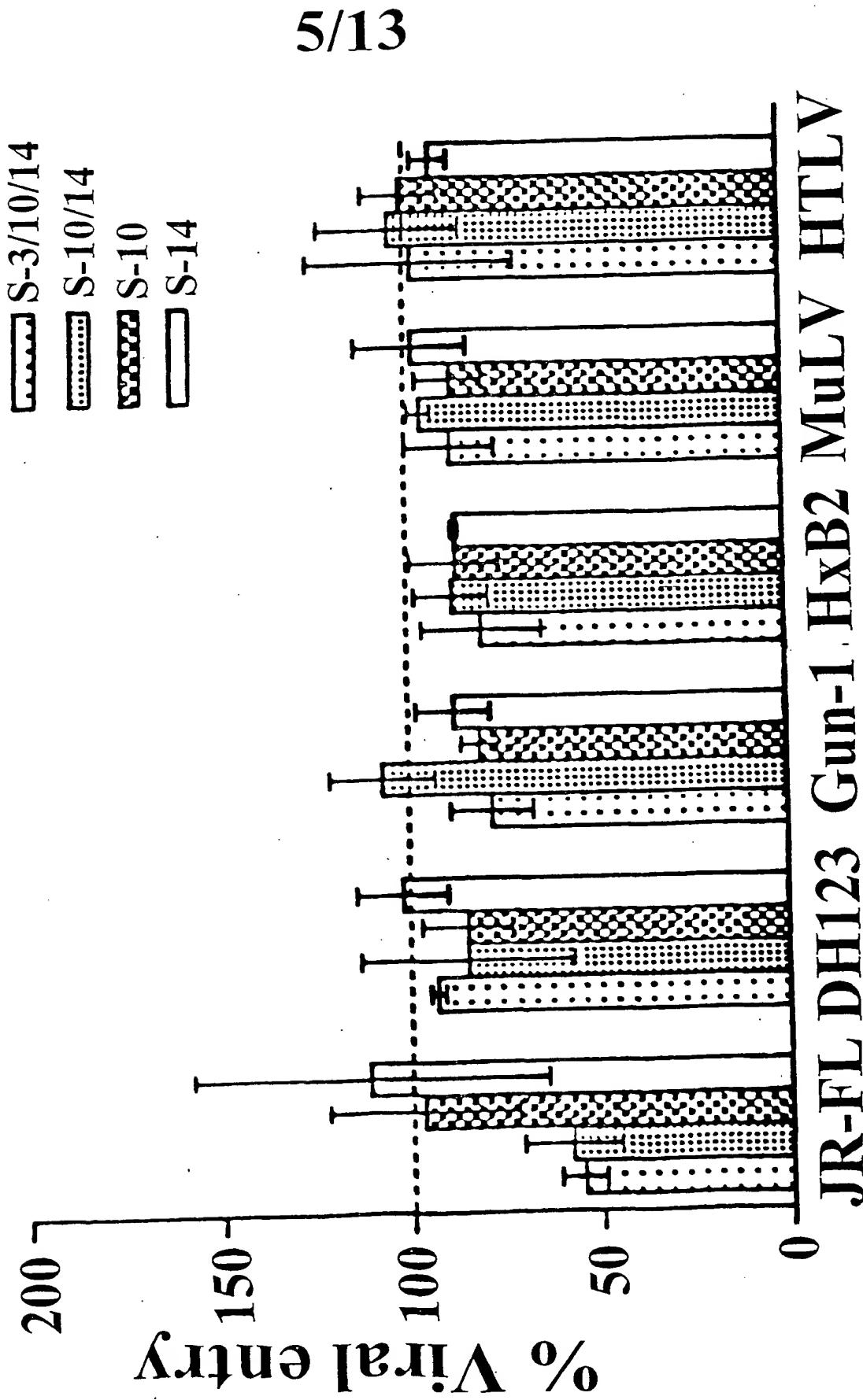


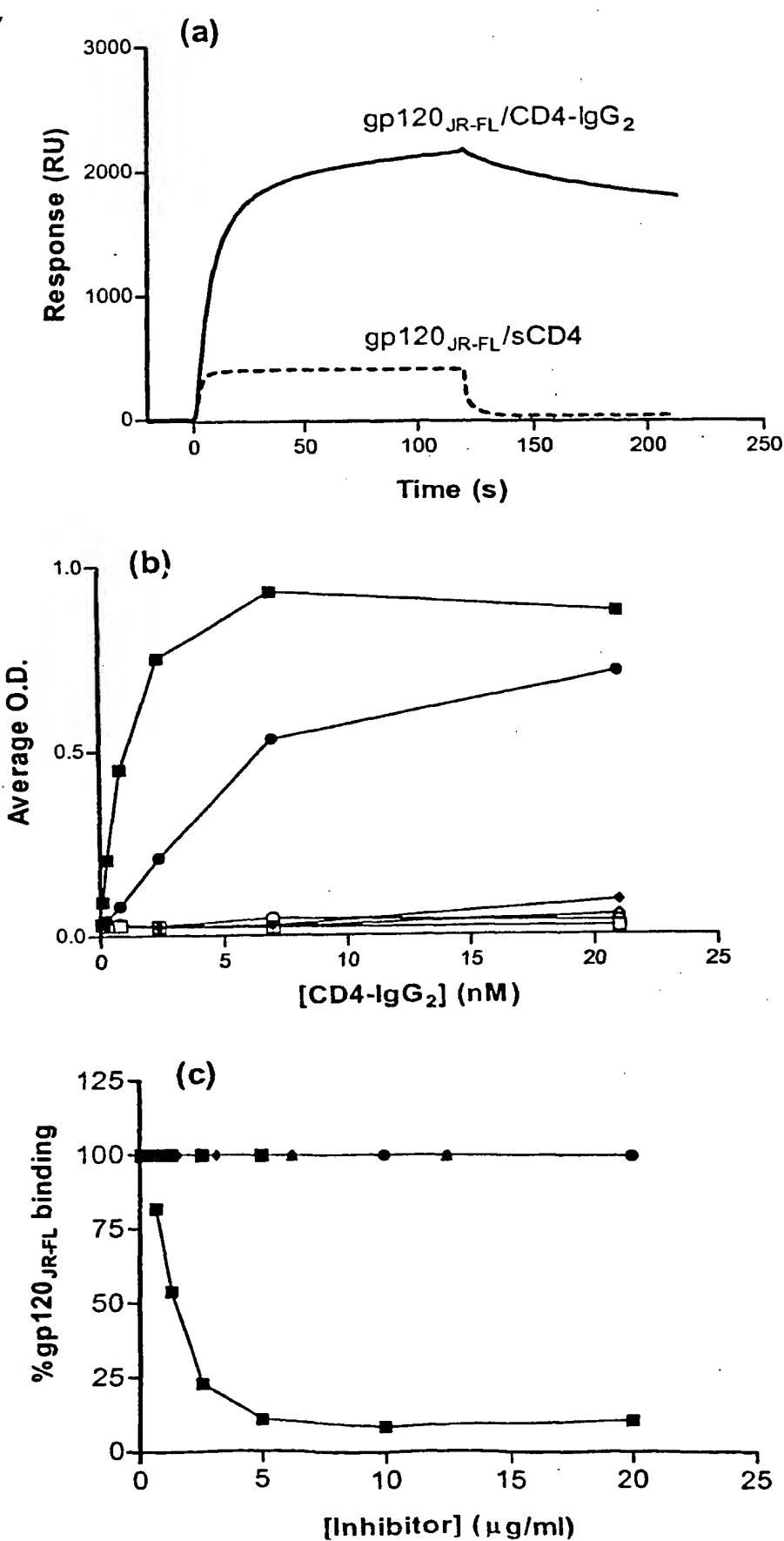
FIGURE 6

| SEQUENCE  | LABEL     |
|---|-----------|
| Unmodified peptide<br>D Y Q V S S P I Y D I N Y Y T S E | 3/10/14   |
| Sulfated peptides                                       |           |
| D <b>Y</b> Q V S S P I <b>Y</b> D I N <b>Y</b> Y T S E  | S-3/10/14 |
| D Y Q V S S P I <b>Y</b> D I N <b>Y</b> Y T S E         | S-10/14   |
| D Y Q V S S P I <b>Y</b> D I N <b>Y</b> Y T S E         | bS-10/14  |
| D Y Q V S S P I <b>Y</b> D I N Y Y T S E G A G K-biotin | S-10      |
| D Y Q V S S P I <b>Y</b> D I N Y Y T S E                | S-14      |
| D Y Q V S S P I <b>Y</b> D I N <b>Y</b> Y T S E         | TS-10/14  |
| Phosphorylated peptides                                 |           |
| D <b>Y</b> Q V S S P I <b>Y</b> D I N <b>Y</b> Y T S E  | P-3/10/14 |
| D Y Q V S S P I <b>Y</b> D I N <b>Y</b> Y T S E         | P-10/14   |
| D Y Q V S S P I <b>Y</b> D I N <b>Y</b> Y T S E         | bP-10/14  |
| Sulfated and Scrambled peptides                         |           |
| <b>Y</b> V S Q P D N T Y I <b>Y</b> S Y E S I D         | SS-2/12   |
| S I D I Y N P T <b>Y</b> V S N <b>Y</b> E S D Y         | SS-10/14  |

6/13

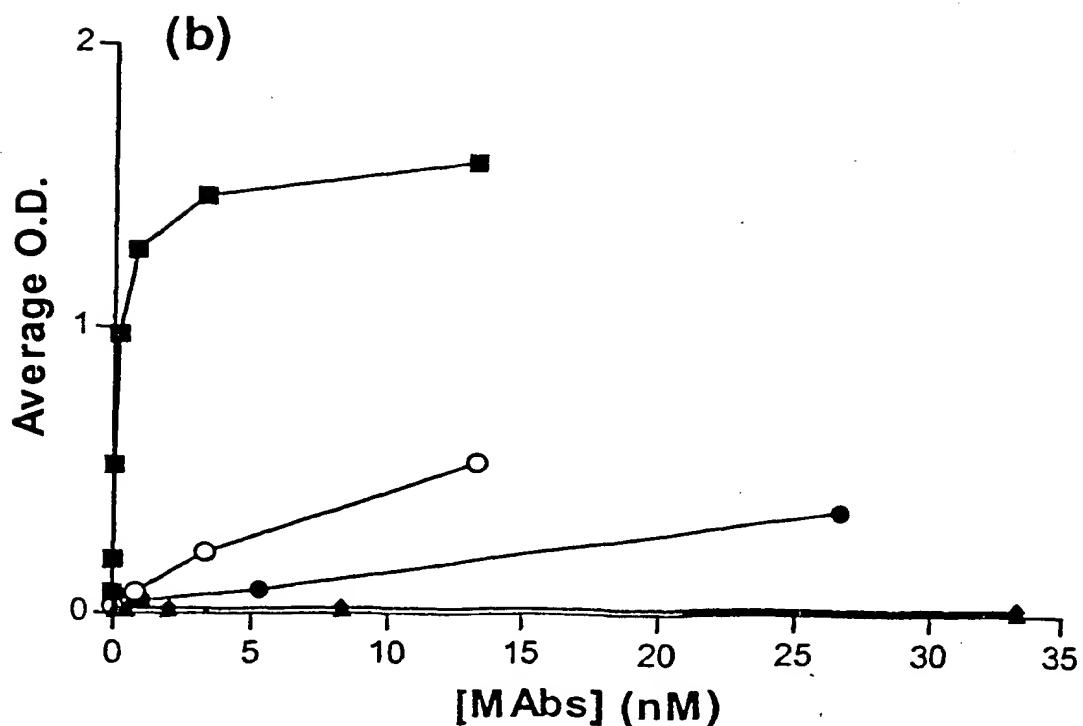
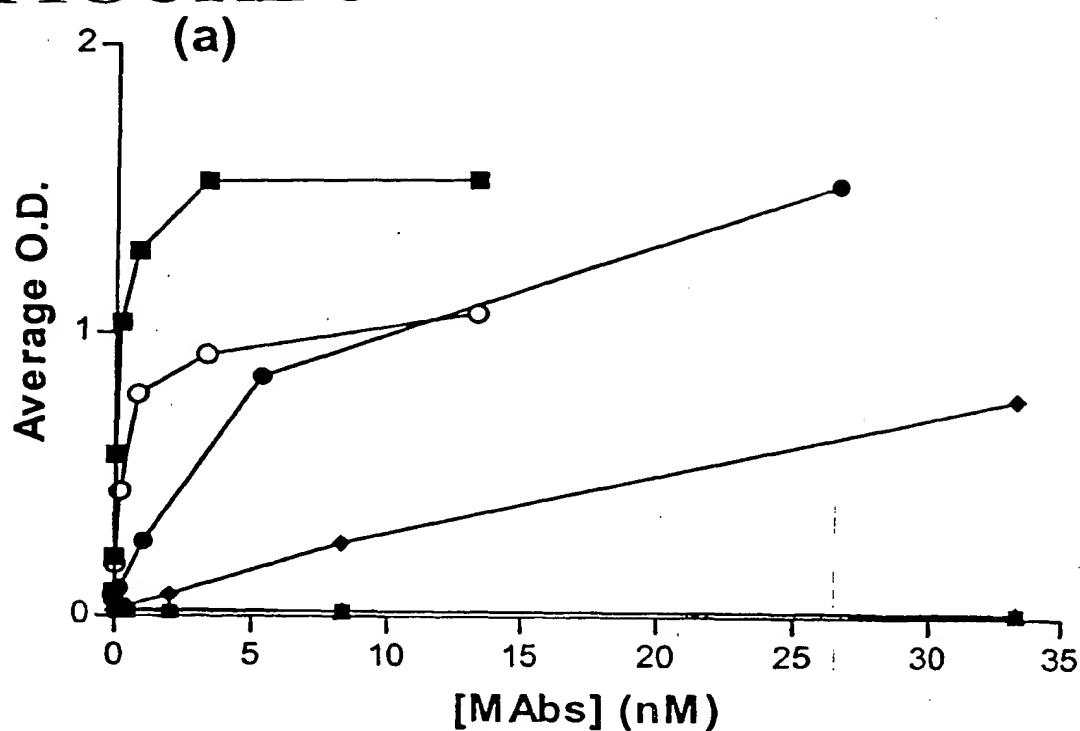
7/13

FIGURE 7



8/13

FIGURE 8



9/13

FIGURE 9

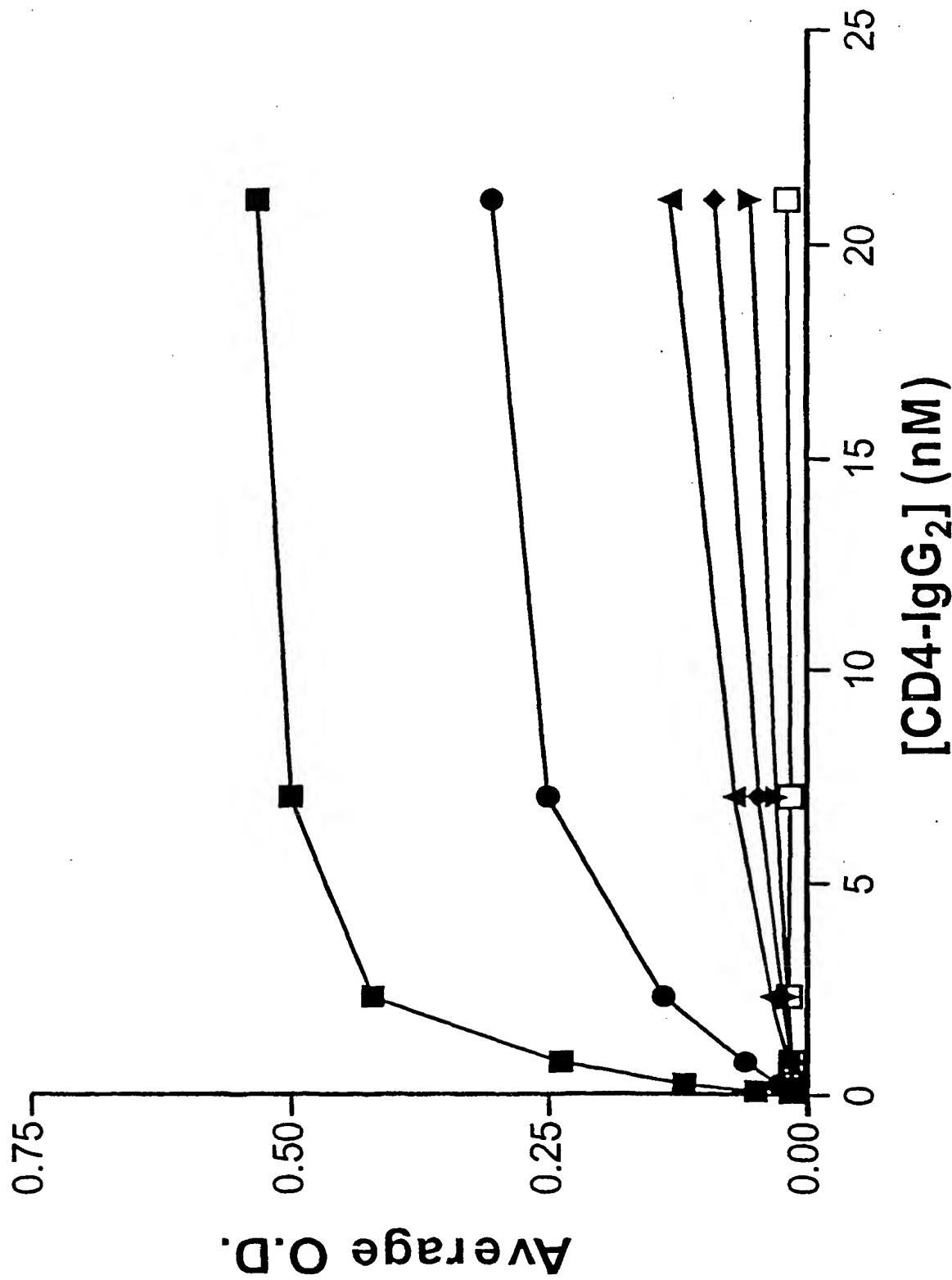


FIGURE 10

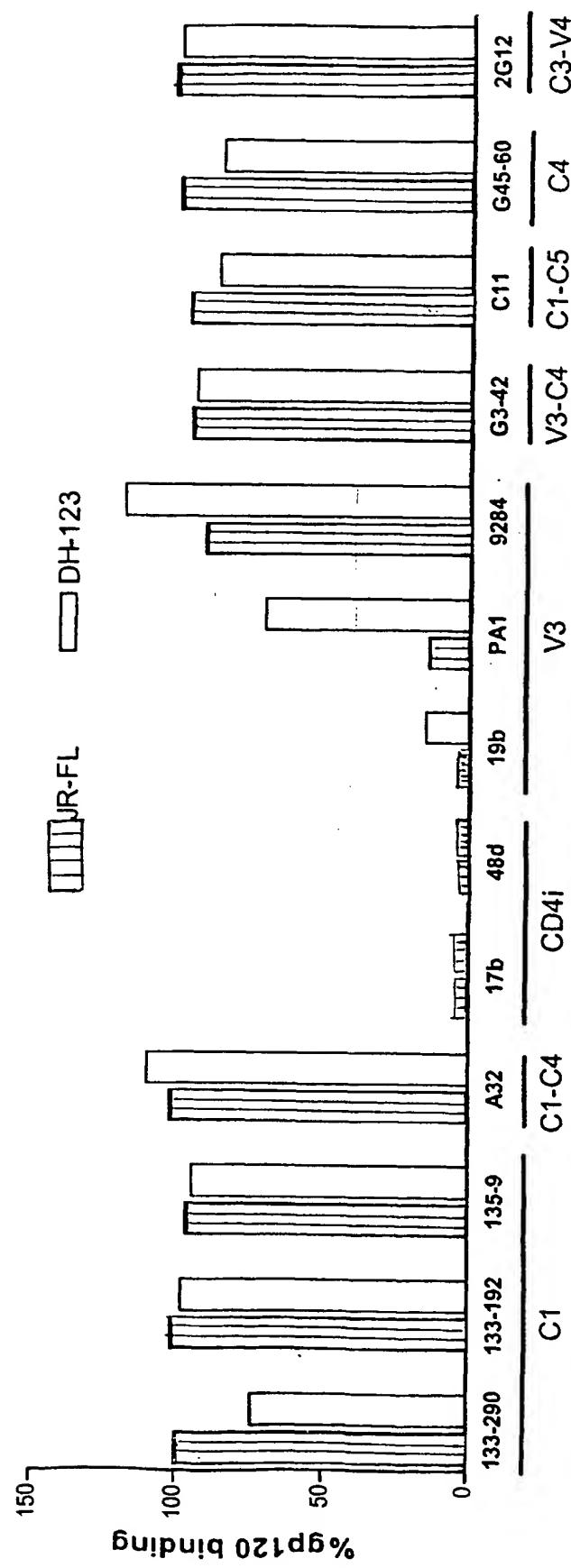
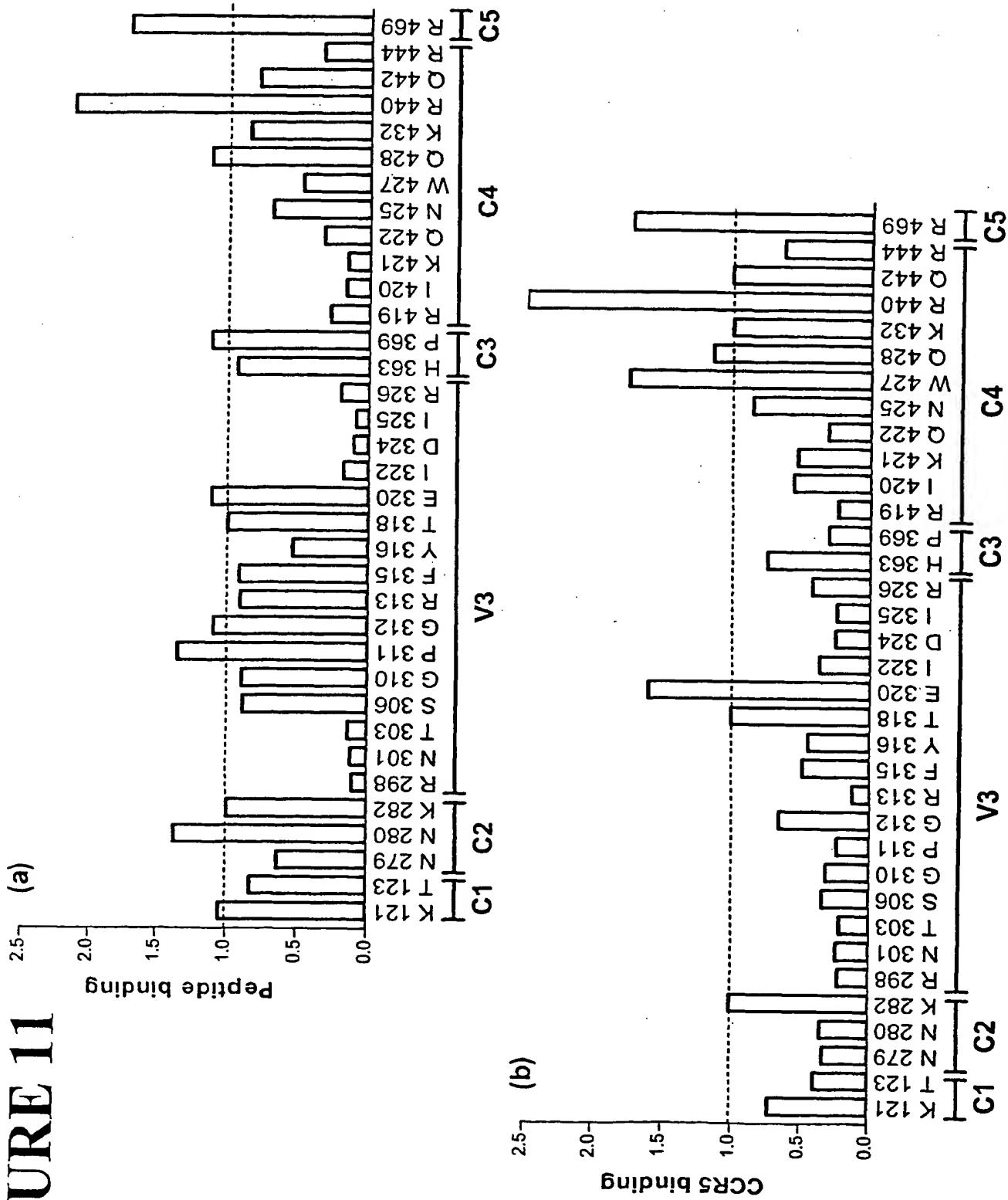


FIGURE 11



12/13

## FIGURE 12

13/13

FIGURE 13

| C1 | C2  | V3  | C3    | C4    | C5    |
|----|-----|-----|-------|-------|-------|
| 96 | 100 | 100 | 96    | 100   | R 469 |
|    |     |     |       |       | R 444 |
|    |     |     |       |       | Q 442 |
|    |     |     |       |       | R 440 |
|    |     |     |       |       | K 432 |
|    |     |     |       |       | Q 428 |
|    |     |     |       |       | W 427 |
|    |     |     |       |       | N 425 |
|    |     |     |       |       | Q 422 |
|    |     |     |       |       | K 421 |
|    |     |     |       |       | I 420 |
|    |     |     |       |       | R 419 |
|    |     |     | P 369 |       |       |
|    |     |     | H 363 |       |       |
|    |     |     |       | R 326 |       |
|    |     |     |       | I 325 |       |
|    |     |     |       | D 324 |       |
|    |     |     |       | I 322 |       |
|    |     |     |       | E 320 |       |
|    |     |     |       | T 318 |       |
|    |     |     |       | Y 316 |       |
|    |     |     |       | F 315 |       |
|    |     |     |       | R 313 |       |
|    |     |     |       | G 312 |       |
|    |     |     |       | P 311 |       |
|    |     |     |       | G 310 |       |
|    |     |     |       | S 306 |       |
|    |     |     |       | T 303 |       |
|    |     |     |       | N 301 |       |
|    |     |     |       | R 298 |       |
|    |     |     |       | K 282 |       |
|    |     |     |       | N 280 |       |
|    |     |     |       | N 279 |       |
|    |     |     |       | T 123 |       |
|    |     |     |       | K 121 |       |

## SEQUENCE LISTING

<110> Progenics Pharmaceuticals, Inc., et al.

<120> SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION

<130> 2048/61010-A-PCT/JPW/SHS/AX

<140> NOT YET KNOWN

<141> 2001-02-28

<160> 17

<170> PatentIn version 3.0

<210> 1

<211> 352

<212> PRT

<213> human

<400> 1

Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr  
1 5 10 15

Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu  
20 25 30

Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn  
35 40 45

Met Leu Val Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met  
50 55 60

Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu  
65 70 75 80

Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe  
85 90 95

Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe  
100 105 110

Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu  
115 120 125

Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe  
130 135 140

Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser  
145 150 155 160

Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr  
165 170 175

Thr Cys Ser Ser His Phe Pro Tyr Ser Gln Tyr Gln Phe Trp Lys Asn  
180 185 190

Phe Gln Thr Leu Lys Ile Val Ile Leu Gly Leu Val Leu Pro Leu Leu  
195 200 205

Val Met Val Ile Cys Tyr Ser Gly Ile Leu Lys Thr Leu Leu Arg Cys  
210 215 220

Arg Asn Glu Lys Lys Arg His Arg Ala Val Arg Leu Ile Phe Thr Ile  
225 230 235 240

Met Ile Val Tyr Phe Leu Phe Trp Ala Pro Tyr Asn Ile Val Leu Leu  
245 250 255

Leu Asn Thr Phe Gln Glu Phe Phe Gly Leu Asn Asn Cys Ser Ser Ser  
260 265 270

Asn Arg Leu Asp Gln Ala Met Gln Val Thr Glu Thr Leu Gly Met Thr  
275 280 285

His Cys Cys Ile Asn Pro Ile Ile Tyr Ala Phe Val Gly Glu Lys Phe  
290 295 300

Arg Asn Tyr Leu Leu Val Phe Phe Gln Lys His Ile Ala Lys Arg Phe  
305 310 315 320

Cys Lys Cys Cys Ser Ile Phe Gln Gln Glu Ala Pro Glu Arg Ala Ser  
325 330 335

Ser Val Tyr Thr Arg Ser Thr Gly Glu Gln Glu Ile Ser Val Gly Leu  
340 345 350

<210> 2

<211> 1376

<212> DNA

<213> human

<400> 2

gaattcccc aacagagcca agctctccat ctagtgaca gggaaagctag cagcaaacct 60  
tcccttcact acaaaaacttc attgcttggc caaaaagaga gttaattcaa ttagacatc 120  
tatgtaggca attaaaaacc tattgtatgta taaaacagt tgcattcatg gagggcaact 180  
aaatacattc taggacttta taaaagatca ctttttattt atgcacaggg tggaaacaaga 240  
tggattatca agtgtcaagt ccaatctatg acatcaatta ttatacatcg gagccctgcc 300  
aaaaaaatcaa tggaaagcaa atcgagccc gcctccgtcc tccgctctac tcactgggt 360  
tcatcttgg ttttgtggc aacatgctgg tcatcctcat cctgataaac tgcaaaaggc 420  
tgaagagcat gactgacatc tacctgctca acctggccat ctctgacctg ttttccttc 480  
ttactgtccc cttctggct cactatgctg ccggccagtg ggactttgga aatacaatgt 540  
gtcaactctt gacagggtct tattttatag gcttcttc tggaaatcttcc ttcatcatcc 600  
tcctgacaat cgataggtac ctggctgtcg tccatgctgt gtttgcttta aaagccagga 660  
cggtcacctt tgggggtggtg acaagtgtga tcacttgggt ggtggctgtg tttgcgtctc 720  
tcccaggaat catcittacc agatcicaaa aagaaggct tcattacacc tgcagctctc 780  
atttccata cagtcagttt caattctggaa agaatttcca gacattaaag atagtcatct 840  
tggggctgggt cctgcccgtg cttgtcatgg tcatctgcta ctgggaatc ctaaaaactc 900

tgcttcggtg tcgaaatgag aagaagaggc acagggctgt gaggcttata tcaccatca 960  
tgattgtta tttctcttc tgggctccct acaacattgt ctttccttg aacacctcc 1020  
aggaattctt tggcctgaat aattgcagta gctctaacag gttggaccaa gctatgcagg 1080  
tgacagagac tcttgggatg acgcactgct gcatcaaccc catcatctat gcctttgtcg 1140  
gggagaagtt cagaaactac ctcttagtct tcttccaaaa gcacattgcc aaacgcttct 1200  
gcaaattgctg ttctattttc cagcaagagg ctcccgagcg agcaagctca gtttacaccc 1260  
gatccactgg ggagcaggaa atatctgtgg gcttgcaca cggactcaag tgggctggtg 1320  
acccagtcag agttgtgcac atggcttagt tttcatacac agcctgggct ggggg 1376

<210> 3

<211> 49

<212> PRT

<213> human immunodeficiency virus

<400> 3

Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg  
1 5 10 15

Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile  
20 25 30

Lys Gln Leu Gln Ala Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp  
35 40 45

Gln

<210> 4

<211> 35

<212> PRT

<213> human immunodeficiency virus

<400> 4

Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser Leu Ile His  
1 5 10 15

Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu  
20 25 30

Leu Leu Glu

35

<210> 5

<211> 22

<212> PRT

<213> human

<400> 5

Leu Leu Thr Val Glu Gln Ala Leu Ala Asp Phe Ala Glu Leu Leu Arg  
1 5 10 15

Ala Leu Arg Arg Asp Leu  
20

<210> 6

<211> 33

<212> PRT

<213> human

<400> 6

His Met Lys Gln Leu Glu Asp Lys Val Glu Glu Leu Leu Ser Lys Asn  
1 5 10 15

Tyr His Leu Glu Asn Glu Val Ala Arg Leu Lys Lys Leu Val Gly Glu  
20 25 30

Arg

<210> 7

<211> 33

<212> PRT

<213> human

<400> 7

His Met Lys Gln Ile Glu Asp Lys Ile Glu Glu Ile Leu Ser Lys Ile  
1 5 10 15

Tyr His Ile Glu Asn Glu Ile Ala Arg Ile Lys Lys Leu Ile Gly Glu  
20 25 30

Val

<210> 8

<211> 40

<212> PRT

<213> human

<400> 8

Leu Thr Asp Thr Leu Gln Ala Glu Thr Asp Gln Leu Glu Asp Glu Lys  
1 5 10 15

Ser Ala Leu Gln Thr Glu Ile Ala Asn Leu Leu Lys Glu Lys Glu Lys  
20 25 30

Leu Glu Phe Ile Leu Ala Ala Arg  
35 40

<210> 9

<211> 40

<212> PRT

<213> human

<400> 9

His Met Arg Arg Ile Ala Arg Leu Glu Glu Lys Val Lys Thr Leu Lys  
1 5 10 15

Ala Gln Asn Ser Glu Leu Ala Ser Thr Ala Asn Met Leu Arg Glu Gln  
20 25 30

Val Ala Gln Leu Lys Gln Lys Tyr  
35 40

<210> 10

<211> 36

<212> PRT

<213> unknown

<220>

<221> PEPTIDE

<222> (1)..(36)

<223> T-20

<400> 10

Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln  
1 5 10 15

Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu  
20 25 30

Trp Asn Trp Phe  
35

<210> 11

<211> 38

<212> PRT

<213> unknown

<220>

<221> PEPTIDE

<222> (1)..(38)

<223> DP107

<400> 11

Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu  
1 5 10 15

Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Ile Leu Ala Val Glu  
20 25 30

Arg Tyr Leu Lys Asp Gln  
35

<210> 12

<211> 34

<212> PRT

<213> unknown

<220>

<221> PEPTIDE

<222> (1)..(34)

<223> N34

<400> 12

Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala  
1 5 10 15

Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln  
20 25 30

Ala Arg

<210> 13

<211> 28

<212> PRT

<213> unknown

<220>

<221> PEPTIDE

<222> (1)..(28)

<223> C28

<400> 13

Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser Leu Ile His  
1 5 10 15

Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys  
20 25

<210> 14

<211> 68

<212> PRT

<213> unknown

<220>

<221> PEPTIDE

<222> (1)..(68)

<223> N34(L6)C28

<400> 14

Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala  
1 5 10 15

Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln  
20 25 30

Ala Arg Ser Gly Gly Arg Gly Gly Trp Met Glu Trp Asp Arg Glu Ile  
35 40 45

Asn Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn  
50 55 60

Gln Gln Glu Lys  
65

<210> 15

<211> 39

<212> PRT

<213> unknown

<220>

<221> PEPTIDE

<222> (1)..(39)

<223> T1249

<400> 15

Trp Gln Glu Trp Glu Gln Lys Ile Thr Ala Leu Leu Glu Gln Ala Gln  
1 5 10 15

Ile Gln Gln Glu Lys Asn Glu Tyr Glu Leu Gln Lys Leu Asp Lys Trp  
 20            25            30

Ala Ser Leu Trp Glu Trp Phe  
 35

<210> 16

<211> 502

<212> PRT

<213> human immunodeficiency virus

<400> 16

Met Arg Val Lys Gly Ile Arg Lys Ser Tyr Gln Tyr Leu Trp Lys Gly  
 1            5            10            15

Gly Thr Leu Leu Leu Gly Ile Leu Met Ile Cys Ser Ala Val Glu Lys  
 20            25            30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr  
 35            40            45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val  
 50            55            60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro  
 65            70            75            80

Gln Glu Val Val Leu Glu Asn Val Thr Glu His Phe Asn Met Trp Lys  
 85            90            95

Asn Asn Met Val Glu Gln Met Gln Glu Asp Ile Ile Ser Leu Trp Asp  
 100            105            110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu  
 115            120            125

Asn Cys Lys Asp Val Asn Ala Thr Asn Thr Asn Asp Ser Glu Gly  
 130            135            140

Thr Met Glu Arg Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr  
145 150 155 160

Ser Ile Arg Asp Glu Val Gln Lys Glu Tyr Ala Leu Phe Tyr Lys Leu  
165 170 175

Asp Val Val Pro Ile Asp Asn Asn Asn Thr Ser Tyr Arg Leu Ile Ser  
180 185 190

Cys Asp Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Ile Ser Phe Glu  
195 200 205

Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys  
210 215 220

Cys Asn Asp Lys Thr Phe Asn Gly Lys Gly Pro Cys Lys Asn Val Ser  
225 230 235 240

Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu  
245 250 255

Leu Leu Asn Gly Ser Leu Ala Glu Glu Val Val Ile Arg Ser Asp  
260 265 270

Asn Phe Thr Asn Asn Ala Lys Thr Ile Ile Val Gln Leu Lys Glu Ser  
275 280 285

Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile  
290 295 300

His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Gly Glu Ile Ile Gly  
305 310 315 320

Asp Ile Arg Gln Ala His Cys Asn Ile Ser Arg Ala Lys Trp Asn Asp  
325 330 335

Thr Leu Lys Gln Ile Val Ile Lys Leu Arg Glu Gln Phe Glu Asn Lys  
340 345 350

Thr Ile Val Phe Asn His Ser Ser Gly Gly Asp Pro Glu Ile Val Met  
355 360 365

His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr Gln  
370 375 380

Leu Phe Asn Ser Thr Trp Asn Asn Asn Thr Glu Gly Ser Asn Asn Thr  
 385            390            395            400

Glu Gly Asn Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn  
 405            410            415

Met Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly  
 420            425            430

Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp  
 435            440            445

Gly Gly Ile Asn Glu Asn Gly Thr Glu Ile Phe Arg Pro Gly Gly Gly  
 450            455            460

Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val  
 465            470            475            480

Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val  
 485            490            495

Val Gln Arg Glu Lys Arg  
 500

<210> 17

<211> 511

<212> PRT

<213> human immunodeficiency virus

<400> 17

Met Arg Val Lys Glu Lys Tyr Gln His Leu Trp Arg Trp Gly Trp Arg  
 1            5            10            15

Trp Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Thr Glu  
 20            25            30

Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala

35                    40                    45

Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu  
50                    55                    60

Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn  
65                    70                    75                    80

Pro Gln Glu Val Val Leu Val Asn Val Thr Glu Asn Phe Asn Met Trp  
85                    90                    95

Lys Asn Asp Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp  
100                    105                    110

Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Ser  
115                    120                    125

Leu Lys Cys Thr Asp Leu Lys Asn Asp Thr Asn Thr Asn Ser Ser Ser  
130                    135                    140

Gly Arg Met Ile Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn  
145                    150                    155                    160

Ile Ser Thr Ser Ile Arg Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe  
165                    170                    175

Tyr Lys Leu Asp Ile Ile Pro Ile Asp Asn Asp Thr Thr Ser Tyr Lys  
180                    185                    190

Leu Thr Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val  
195                    200                    205

Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala  
210                    215                    220

Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr  
225                    230                    235                    240

Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser  
245                    250                    255

Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Val Val Ile  
260                    265                    270

Arg Ser Val Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu  
275 280 285

Asn Thr Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg  
290 295 300

Lys Arg Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala Phe Val Thr Ile  
305 310 315 320

Gly Lys Ile Gly Asn Met Arg Gln Ala His Cys Asn Ile Ser Arg Ala  
325 330 335

Lys Trp Asn Asn Thr Leu Lys Gln Ile Ala Ser Lys Leu Arg Glu Gln  
340 345 350

Phe Gly Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser Ser Gly Gly Asp  
355 360 365

Pro Glu Ile Val Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr  
370 375 380

Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp  
385 390 395 400

Ser Thr Glu Gly Ser Asn Asn Thr Glu Gly Ser Asp Thr Ile Thr Leu  
405 410 415

Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Lys Val Gly Lys  
420 425 430

Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser Ser Asn  
435 440 445

Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Ser Asn Asn Glu  
450 455 460

Ser Glu Ile Phe Arg Pro Gly Gly Asp Met Arg Asp Asn Trp Arg  
465 470 475 480

Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val  
485 490 495

Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg  
500 505 510